There is No EPTAS for Two-dimensional Knapsack

Ariel Kulik* Hadas Shachnai†

Abstract

In the *d-dimensional (vector) knapsack* problem given is a set of items, each having a d-dimensional size vector and a profit, and a d-dimensional bin. The goal is to select a subset of the items of maximum total profit such that the sum of all vectors is bounded by the bin capacity in each dimension. It is well known that, unless $P = NP$, there is no *fully polynomial time approximation scheme* for d-dimensional knapsack, already for $d = 2$. The best known result is a *polynomial time approximation scheme (PTAS)* due to Frieze and Clarke (*European J. of Operational Research, 100–109, 1984*) for the case where $d \geq 2$ is some fixed constant. A fundamental open question is whether the problem admits an *efficient PTAS (EPTAS)*.

In this paper we resolve this question by showing that there is no EPTAS for d-dimensional knapsack, already for $d = 2$, unless $W[1] = FPT$. Furthermore, we show that unless all problems in SNP are solvable in sub-exponential time, there is no approximation scheme for two-dimensional knapsack whose running time is $f(1/\varepsilon)|\mathcal{I}|^{o(\sqrt{1/\varepsilon})}$, for any function f. Together, the two results suggest that a significant improvement over the running time of the scheme of Frieze and Clarke is unlikely to exist.

Keywords: two-dimensional knapsack, efficient polynomial time approximation schemes, parameterized complexity, theory of computation

1 Introduction

In the well known *d-dimensional knapsack* problem, given is a set of n items $\{1, \ldots, n\}$, where each item i has a d-dimensional size vector $\bar{s}_i \geq 0$, and a profit $p_i > 0$. Also, given is a d-dimensional bin whose capacity is $\bar{B} = (B_1, \ldots, B_d)$. A feasible solution is a subset of the items $A' \subseteq A$ such that the total size of the items in A' in each dimension r is bounded by B_r, $1 \leq r \leq d$. The objective is to find a feasible solution of maximum total profit. The special case where $d = 1$ is the classic *0-1 knapsack* problem.

This paper studies the efficiency of finding $(1 - \varepsilon)$-approximations for d-dimensional knapsack. A maximization problem Π admits a *polynomial-time approximation scheme (PTAS)* if there is an algorithm $\mathcal{A}(\mathcal{I}, \varepsilon)$ such that, for any $\varepsilon > 0$ and any instance $\mathcal{I}$ of Π, $\mathcal{A}(\mathcal{I}, \varepsilon)$ outputs a $(1 - \varepsilon)$-approximate solution in time $|\mathcal{I}|^{f(1/\varepsilon)}$ for some function f. As ε gets smaller, the exponent of the polynomial $|\mathcal{I}|^{f(1/\varepsilon)}$ may become very large. Two important restricted classes of approximation schemes were defined to eliminate this dependence. An *efficient polynomial-time approximation scheme (EPTAS)* is a PTAS whose running time is $f(1/\varepsilon)|\mathcal{I}|^{O(1)}$, whereas a *fully polynomial time approximation scheme (FPTAS)* runs in time $(1/\varepsilon)^{O(1)}|\mathcal{I}|^{O(1)}$.

*Computer Science Dept., Technion, Haifa 32000, Israel. E-mail: kulik@cs.technion.ac.il.

†Computer Science Dept., Technion, Haifa 32000, Israel. E-mail: hadas@cs.technion.ac.il. Work supported by the Technion V.P.R. Fund.

While the classic 0-1 knapsack problem admits an FPTAS, i.e., for any $\varepsilon > 0$, a $(1-\varepsilon)$-approximation for the optimal solution can be found in $O(n/\varepsilon^2 \cdot \log(1/\varepsilon))$ steps [10, 11],[1] packing in higher dimensions (also known as *d-dimensional vector packing*) is substantially harder to solve, exactly or approximately. It is well known that, unless $P = NP$, there is no FPTAS for d-dimensional knapsack, already for $d = 2$ [12, 14] (see also [13],[7]). Frieze and Clarke developed in [6] the first PTAS for the d-dimensional knapsack. Subsequently, a scheme with improved running time of $O(n^{\lceil d/\varepsilon \rceil - d})$ was given by Caprara et al. [1].

As d-dimensional knapsack does not admit an FPTAS, a fundamental open question is whether there exists an EPTAS. In this paper we resolve this question by showing that there is no EPTAS for two-dimensional knapsack, unless $W[1] = FPT$.[2] Furthermore, we use the results of [2] to show that unless all problems in SNP are solvable in sub-exponential time,[3] there is no approximation scheme for two-dimensional knapsack whose running time is $f(1/\varepsilon)|\mathcal{I}|^{o(\sqrt{1/\varepsilon})}$, for any function f. Together, the two results suggest that a significant improvement over the running time of the scheme of [1] is unlikely to exist. We note that, for the case where $d = 1$ an EPTAS exists also for the *multiple knapsack* problem (see the recent work of Jansen [9]).

2 Hardness Results

Denote by $OPT(\mathcal{I})$ the value of an optimal solution for an instance $\mathcal{I}$ of the d-dimensional knapsack problem. We use in the proof of hardness the following parameterized version of the *subset sum* problem, known as *sized subset sum*. Given a set of positive integers $L = \{x_1, \ldots, x_n\}$, and the positive integer S, k, decide if there is a subset $L' \subseteq L$ of size k, such that the sum of elements in L' is *exactly* S (in this case we say that the input is *satisfied*). The sized subset sum problem is known to be $W[1]$-hard [4].

We give a reduction from an instance (L, S, k) of sized subset sum to an instance of two-dimensional knapsack, denoted by $R(L, S, k)$.

Given an instance (L, S, k), we first modify the values of the elements in L. Define

$$\tilde{x}_i = \frac{x_i + \frac{k-1}{k} \cdot S}{k},$$

and let $\tilde{L} = \{\tilde{x}_1, \ldots, \tilde{x}_n\}$. Note that, for any $1 \le i \le n$, $0 \le \tilde{x}_i \le \frac{2 \cdot S}{k}$ (w.l.o.g. $x_i \le S$). An important property of the above transformation is that it does not affect the satisfiability of the original instance.

Lemma 1 *The instance (L, S, k) is satisfied if and only if $(\tilde{L}, S, k)$ is satisfied.*

Proof: If (L, S, k) is satisfied then there is a subset $\{x_{i_1}, \ldots, x_{i_k}\} = L' \subseteq L$ such that $\sum_{j=1}^{k} x_{i_j} = S$. Consider the subset $\{\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_k}\} = \tilde{L}' \subseteq \tilde{L}$, then

$$\sum_{j=1}^{k} \tilde{x}_{i_j} = \sum_{j=1}^{k} \frac{x_{i_j} + \frac{k-1}{k} \cdot S}{k} = \frac{1}{k} \sum_{j=1}^{k} x_{i_j} + \frac{1}{k} \sum_{j=1}^{k} \frac{k-1}{k} \cdot S = S,$$

[1] See also the comprehensive survey of known results in [12].

[2] For the recent theory of fixed-parameter algorithms and parameterized complexity, see, e.g., [5, 3].

[3] The complexity class of SNP was introduced in [15]. The class includes such NP-hard problems as *vertex cover*, *independent set* and *3SAT*, among others. Based on known results in complexity theory, it is unlikely that all of the problems in this class can be solved in sub-exponential time (see [2] and the references therein).

and we have that $(\tilde{L}, S, k)$ is also satisfied.

If $(\tilde{L}, S, k)$ is satisfied, then there is a subset $\{\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_k}\} = \tilde{L}' \subseteq \tilde{L}$ such that $\sum_{j=1}^{k} \tilde{x}_{i_j} = S$. By the definition of $\tilde{L}$, we have that

$$S = \sum_{j=1}^{k} \tilde{x}_{i_j} = \sum_{j=1}^{k} \frac{x_{i_j} + \frac{k-1}{k} \cdot S}{k} = \frac{1}{k} \sum_{j=1}^{k} x_{i_j} + \frac{1}{k} \sum_{j=1}^{k} \frac{k-1}{k} \cdot S = \frac{1}{k} \sum_{j=1}^{k} x_{i_j} + \frac{k-1}{k} \cdot S,$$

and $\sum_{j=1}^{k} x_{i_j} = S$. Thus, (L, S, k) is satisfied as well. $\square$

Now, we define the instance $R(L, S, k)$ of two-dimensional knapsack. The items are $\{1, \ldots, n\}$, where each item i has size $\bar{s}_i = (\tilde{x}_i, \frac{2 \cdot S}{k} - \tilde{x}_i)$ and unit profit. Let $s_{i,1}$ and $s_{i,2}$ denote the first and second entries to the vector $\bar{s}_i$, respectively. The capacity of the bin is $\bar{B} = (S, S)$. Note that $R(L, S, k)$ can be computed in polynomial time in the size of the instance (L, S, k), and its size is also polynomial.

Lemma 2 *$OPT(R(L, S, k)) \leq k$.*

Proof: Assume that there is a feasible subset of items $A \subseteq \{1, \ldots, n\}$ whose value is greater than k for $R(L, S, k)$, then $|A| \geq k+1$. Since A is feasible, we have that $\sum_{i \in A} s_{i,1} = \sum_{i \in A} \tilde{x}_i \leq S$, and thus

$$S \geq \sum_{i \in A} s_{i,2} = \sum_{i \in A} \left(\frac{2 \cdot S}{k} - \tilde{x}_i \right) = |A| \cdot \frac{2 \cdot S}{k} - \sum_{i \in A} \tilde{x}_i > S,$$

a contradiction. $\square$

Lemma 3 *The instance $(\tilde{L}, S, k)$ is satisfied if and only if $OPT(R(L, S, k)) \geq k$.*

Proof: If the instance $(\tilde{L}, S, k)$ is satisfied then there is a subset $\{\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_k}\} = \tilde{L}' \subseteq \tilde{L}$ such that $\sum_{j=1}^{k} \tilde{x}_{i_j} = S$. Thus, the solution $A = \{i_1, \ldots, i_k\}$ for $R(L, S, k)$ is feasible in both dimensions, i.e., $\sum_{j=1}^{k} s_{i_j,1} = \sum_{j=1}^{k} \tilde{x}_{i_j} = S$, and also $\sum_{j=1}^{k} s_{i_j,2} = \sum_{j=1}^{k} \left(\frac{2 \cdot S}{k} - \tilde{x}_{i_j} \right) = S$. The value of this solution is k, therefore $OPT(R(L, S, k)) \geq k$.

If $OPT(R(L, S, k)) \geq k$ then, by Lemma 2, we have that $OPT(R(L, S, k)) = k$. Let $A = \{i_1, \ldots, i_k\}$ be an optimal solution, then

$$S \geq \sum_{j=1}^{k} s_{i_j,2} = \sum_{j=1}^{k} \left(\frac{2 \cdot S}{k} - \tilde{x}_{i_j} \right) = 2S - \sum_{j=1}^{k} \tilde{x}_{i_j},$$

and we have that $\sum_{j=1}^{k} \tilde{x}_{i_j} \geq S$. On the other hand, $S \geq \sum_{j=1}^{k} s_{i_j,1} = \sum_{j=1}^{k} \tilde{x}_{i_j}$, and thus $\sum_{j=1}^{k} \tilde{x}_{i_j} = S$. It follows that $(\tilde{L}, S, k)$ is satisfied. $\square$

By the above discussion, we have the next lemma.

Lemma 4 *For any instance (L, S, k) of sized subset sum, (L, S, k) is satisfied if and only if $OPT(R(L, S, k)) \geq k$.*

Proof: The statement of the lemma follows immediately from Lemmas 1 and 3. $\square$

Suppose that we have an approximation scheme $\mathcal{A}(\mathcal{I}, \varepsilon)$ for two-dimensional knapsack. We now show how $\mathcal{A}$ can be used to decide if an input for sized subset sum is satisfied.

Lemma 5 *Let $\mathcal{A}(\mathcal{I}, \varepsilon)$ be an approximation scheme for two-dimensional knapsack with running time $f(1/\varepsilon) \cdot |\mathcal{I}|^{g(1/\varepsilon)}$, then there is an algorithm for sized subset sum with running time $f(2k) \cdot |(L, S, k)|^{O(g(2k))}$.*

Proof: Consider the following algorithm for sized subset sum. Given an instance (L, S, k), define the input for two-dimensional knapsack $\mathcal{I} = R(L, S, k)$, and run $\mathcal{A}(\mathcal{I}, \frac{1}{2k})$. If the optimal solution output by the algorithm is of value at least k return that (L, S, k) is satisfied, otherwise return that it cannot be satisfied.

Note that if $OPT(\mathcal{I}) \geq k$, the value output by $\mathcal{A}$ is at least $\left(1 - \frac{1}{2k}\right) k = k - \frac{1}{2} > k - 1$. On the other hand, if $OPT(\mathcal{I}) < k$, the output value is at most $k - 1$. Also, by Lemma 4, (L, S, k) is satisfied if and only if $OPT(\mathcal{I}) \geq k$. Hence, the algorithm decides correctly if (L, S, k) is satisfied.

The construction of $\mathcal{I}$ takes polynomial time in $|(L, S, k)|$, and running $\mathcal{A}$ on the instance $\mathcal{I}$ requires $f(2k) \cdot |R(L, S, k)|^{O(g(2k))}$ steps. Thus, the running time of the algorithm is $f(2k) \cdot |(L, S, k)|^{O(g(2k))}$. □

We summarize in our main result.

Theorem 6 *There is no EPTAS for two-dimensional knapsack unless $W[1] = FPT$.*

Proof: Assume there is an $EPTAS$ for two-dimensional knapsack. That is, there exists an algorithm $\mathcal{A}(\mathcal{I}, \varepsilon)$ that, given an instance $\mathcal{I}$ for the problem, returns a $(1 - \varepsilon)$-approximation for the optimal solution in $f(1/\varepsilon) \cdot |\mathcal{I}|^c$ steps. Then, by Lemma 5, there is an algorithm for sized subset sum whose running time is $f(2k) \cdot |(L, S, k)|^{c'}$. It follows that sized subset sum is *fixed parameter tractable*, which cannot hold unless $W[1] = FPT$. □

The standard parametrization of two-dimensional knapsack is as follows. Given an instance of the problem in which all values are integral, and an integer $k \geq 1$, decide if there is a feasible solution of value k or greater. In fact, we have shown the following.

Theorem 7 *The standard parametrization of two-dimensional knapsack is $W[1]$-hard.*

We can use the same reduction to derive an explicit lower bound on the running time of approximation schemes for two-dimensional knapsack, under a different complexity measure. To do so, we first derive a lower bound on the complexity of sized subset sum.

Chen at el. show in [2] that unless all problems in SNP are solvable in sub-exponential time, there is no algorithm for *independent set* whose running time is $f(k)m^{o(k)}$, where m is the input length. Downey and Fellows [4] give a reduction from *independent set* to *perfect code* in which, given a graph G and a parameter k, a new graph H is constructed, such that G has an independent set of size k iff H has a perfect code of size $k' = \frac{k(k+1)}{2} + k + 1$. Under the same assumption, this implies that there is no algorithm for perfect code with running time $f(k)m^{o(\sqrt{k})}$, where m is the input size. Furthermore, a reduction given in [4], from perfect code with a parameter k to sized subset sum with the same parameter k, implies that there is no algorithm for sized subset sum with running time $f(k) \cdot |\mathcal{I}|^{o(\sqrt{k})}$. This is summarized in the next result.

Lemma 8 *Unless all problems in SNP are solvable in sub-exponential time, there is no algorithm for sized subset sum whose running time is $f(k) \cdot |\mathcal{I}|^{o(\sqrt{k})}$, for any function f, where $|\mathcal{I}|$ is the input size.*

From the above discussion, we have

Theorem 9 *Unless all problems in SNP are solvable in sub-exponential time, there is no approximation scheme for two-dimensional knapsack with running time* $f(1/\varepsilon)|\mathcal{I}|^{o(\sqrt{1/\varepsilon})}$*, for any function* f*, where* $|\mathcal{I}|$ *is the size of the input for the problem.*

Proof: Assume that there is an approximation scheme $\mathcal{A}(\mathcal{I}, \varepsilon)$ for two-dimensional knapsack with running time $f(1/\varepsilon)|\mathcal{I}|^{o(\sqrt{1/\varepsilon})}$, for some function f. Thus, by Lemma 5, there is an algorithm for sized subset sum whose running time is $f(2k)|\mathcal{I}|^{o(\sqrt{k})}$. By Lemma 8, this cannot hold unless all problems in SNP are solvable in sub-exponential time. □

In conclusion, we comment that our reductions yield a restricted class of highly structured inputs for d-dimensional knapsack, which may not reflect the set of inputs arising in real-life applications. For many inputs, it seems reasonable to assume that a small modification in the bin capacity would result in a small change in the profit of an optimal solution for the given instance. For such inputs, *augmenting algorithms*, i.e., algorithms that output a solution with profit at least as high as the optimal, while violating the bin capacity (in any dimension) at most by factor $(1 + \varepsilon)$, seem to fit well. For fixed values of d, an augmenting algorithm, with running time polynomial in $1/\varepsilon$ and in the input size, can be used to obtain a feasible solution whose profit is at least $1 - \varepsilon$ of the optimal. [4]

References

[1] A. Caprara, H. Kellerer, U. Pferschy, and D. Pisinger. Approximation algorithms for knapsack problems with cardinality constraints. *European Journal of Operational Research*, 123(2):333 – 345, 2000.

[2] J. Chen, X. Huang, I. A. Kanj, and G. Xia. Linear FPT reductions and computational lower bounds. In *STOC '04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 212–221, New York, NY, USA, 2004. ACM.

[3] R. G. Downey and M. Fellows. *Parameterized Complexity.* Monographs in Computer Science. Springer, 1999.

[4] R. G. Downey and M. R. Fellows. Fixed-parameter tractability and completeness II: On completeness for W[1]. *Theoretical Computer Science*, 141(1-2):109 – 131, 1995.

[5] J. Flum and M. Grohe. *Parameterized Complexity Theory (Texts in Theoretical Computer Science. An EATCS Series).* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[6] A. M. Frieze and M. Clarke. Approximation algorithms for the m-dimensional 0-1 knapsack problem: worst-case and probabilistic analyses. *European J. of Operational Research*, 15(1):100–109, 1984.

[7] G. Gens and E. Levner. Complexity of approximation algorithms for combinatorial problems: a survey. *SIGACT News*, 12(3):52–65, 1980.

[4] Such an algorithm can be obtained by discretizing the item sizes in each dimension, r, to be integral multiples of $\frac{\varepsilon}{n} \cdot B_r$, and using dynamic programming over the maximal profit attainable for each of the possible size vectors. Detailed expositions of these standard techniques are given, e.g., in [8, 16].

[8] D. S. Hochbaum, editor. *Approximation algorithms for NP-hard problems.* PWS Publishing Co. Boston, MA, USA, 1996.

[9] K. Jansen. Parameterized approximation scheme for the multiple knapsack problem. In C. Mathieu, editor, *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 665–674. SIAM, 2009.

[10] H. Kellerer and U. Pferschy. A new fully polynomial approximation scheme for the knapsack problem. *J. Combinatorial Optimization*, 3:59–71, 1999.

[11] H. Kellerer and U. Pferschy. Improved dynamic programming in connection with an FPTAS for the knapsack problem. *J. Combinatorial Optimization*, 8(1):5–11, 2004.

[12] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems.* Springer, October 2004.

[13] B. Korte and R. Schrader. On the existence of fast approximation schemes. *Nonlinear Programming*, 4:415–437, 1981.

[14] M. J. Magazine and M.-S. Chern. A note on approximation schemes for multidimensional knapsack problems. *Mathematics of Operations Research*, 9(2):244–247, 1984.

[15] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. Syst. Sci.*, 43(3):425–440, 1991.

[16] V. Vazirani. *Approximation Algorithms.* Springer Verlag, 2001.