Improved Approximation for Two-Dimensional Vector Multiple Knapsack

Tomer Cohen
Computer Science Department, Technion, Haifa, Israel

Ariel Kulik
CISPA Helmholtz Center for Information Security, Saarbrücken, Germany

Hadas Shachnai
Computer Science Department, Technion, Haifa, Israel

Abstract

We study the UNIFORM 2-DIMENSIONAL VECTOR MULTIPLE KNAPSACK (2VMK) problem, a natural variant of MULTIPLE KNAPSACK arising in real-world applications such as virtual machine placement. The input for 2VMK is a set of items, each associated with a 2-dimensional *weight* vector and a positive *profit*, along with m 2-dimensional bins of uniform (unit) capacity in each dimension. The goal is to find an assignment of a subset of the items to the bins, such that the total weight of items assigned to a single bin is at most one in each dimension, and the total profit is maximized.

Our main result is a $(1 - \frac{\ln 2}{2} - \varepsilon)$-approximation algorithm for 2VMK, for every fixed $\varepsilon > 0$, thus improving the best known ratio of $(1 - \frac{1}{e} - \varepsilon)$ which follows as a special case from a result of [Fleischer at al., MOR 2011]. Our algorithm relies on an adaptation of the Round&Approx framework of [Bansal et al., SICOMP 2010], originally designed for set covering problems, to maximization problems. The algorithm uses randomized rounding of a configuration-LP solution to assign items to $\approx m \cdot \ln 2 \approx 0.693 \cdot m$ of the bins, followed by a reduction to the (1-dimensional) MULTIPLE KNAPSACK problem for assigning items to the remaining bins.

2012 ACM Subject Classification Theory of computation → Packing and covering problems

Keywords and phrases vector multiple knapsack, two-dimensional packing, randomized rounding, approximation algorithms

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2023.20

Related Version *Full Version*: https://arxiv.org/abs/2307.02137

Funding *Ariel Kulik*: Research supported by the European Reseach Concil (ERC) consolidator grant no. 725978 SYSTEMATICGRAPH.

1 Introduction

The KNAPSACK problem and its variants have attracted much attention in the past four decades, and have been instrumental in the development of approximation algorithms. In this paper we study a variant of KNAPSACK which uses components of two well studied knapsack problems: MULTIPLE KNAPSACK and 2-DIMENSIONAL KNAPSACK.

An instance of uniform MULTIPLE KNAPSACK consists of a set I of items of non-negative profits and weights in $[0, 1]$, as well as m uniform (unit size) bins. We seek a subset of the items of maximal total profit which can be packed in the m bins. An instance of 2-DIMENSIONAL KNAPSACK is a set I of items, each has a 2-dimensional weight in $[0, 1]^2$, and a non-negative profit. The objective is to find a subset of the items whose total weight is at most one in each dimension, such that the total profit is maximized. We study a variant of uniform MULTIPLE KNAPSACK where each bin is a 2-dimensional knapsack, thus generalizing both problems.

Formally, an instance of UNIFORM 2-DIMENSIONAL MULTIPLE KNAPSACK (2VMK) is a tuple $\mathcal{I} = (I, w, p, m)$, where I is a set of items, $w : I \to [0,1]^2$ is a 2-dimensional weight function, $p : I \to \mathbb{R}_{\geq 0}$ is a profit function, and m is the number of bins. For any $k \in \mathbb{N}$, let $[k] = \{1, 2, \ldots, k\}$. A *solution* for the instance (I, w, p, m) is a collection of subsets of items $S_1, \ldots, S_m \subseteq I$ such that $w(S_b) = \sum_{i \in S_b} w(i) \leq (1,1)$ for all $b \in [m]$.[1] Our objective is to find a solution $S_1, \ldots, S_m$ which maximizes the total profit, given by $p\left(\bigcup_{b \in [m]} S_b\right)$.

A natural application of 2VMK arises in the cloud computing environment. Consider a data center consisting of m hosts (physical machines). Each host has an available amount of processing power (CPU) and limited memory. For simplicity, these amounts can be scaled to one unit. The data center administrator has a queue of client requests to assign *virtual machines* (VMs) to the hosts.[2] Each VM has a demand for processing power and memory, and its execution is associated with some profit. The administrator needs to assign a subset of the VMs to the hosts, such that the total processing power and memory demands on each host do not exceed the available amounts, and the profit gained from the VMs is maximized (see [5] for other optimization objectives in this setting). Another application of 2VMK comes from spectrum allocation in cognitive radio networks [18].

Our goal is to develop an efficient polynomial-time approximation algorithm for 2VMK. Let $\alpha \in (0, 1]$ be a constant. An algorithm $\mathcal{A}$ is an α-approximation algorithm for 2VMK if for any instance $\mathcal{I}$ of 2VMK it returns in polynomial-time a solution of profit at least $\alpha \cdot \mathrm{OPT}(\mathcal{I})$, where $\mathrm{OPT}(\mathcal{I})$ is the optimal profit for $\mathcal{I}$. A *polynomial-time approximation scheme* (PTAS) is an infinite family $\{A_\varepsilon\}$ of $(1-\varepsilon)$-approximation algorithms, one for each $\varepsilon > 0$. A weaker notion is that of a *randomized* α-approximation algorithm, where the algorithm always returns a solution, but the profit is at least $\alpha \cdot \mathrm{OPT}(\mathcal{I})$ with some constant probability.

As the classic MULTIPLE KNAPSACK problem admits an *efficient PTAS* (EPTAS), even for instances with arbitrary bin capacities [10, 11], and the single bin problem, i.e., 2-DIMENSIONAL KNAPSACK has a PTAS [9], a natural question is whether 2VMK admits a PTAS as well. By a simple reduction from 2-DIMENSIONAL VECTOR BIN PACKING (2VBP), we show that such a PTAS is unlikely to exist.[3]

▶ **Theorem 1.** *Assuming* $\mathrm{P} \neq \mathrm{NP}$ *there is no PTAS for* 2VMK*.*

Thus, we focus on deriving the best constant factor approximation for the problem. For any $\varepsilon > 0$, a randomized $(1 - e^{-1} - \varepsilon) \approx 0.632$-approximation algorithm for 2VMK follows from a result of [8], as a special case of the SEPARABLE ASSIGNMENT PROBLEM (SAP). Our main result is an improved approximation ratio for the problem.

▶ **Theorem 2.** *For every fixed $\varepsilon > 0$, there is a randomized $\left(1 - \frac{\ln 2}{2} - \varepsilon\right) \approx 0.653$-approximation algorithm for* 2VMK*.*

1.1 Prior Work

The special case of 2VMK with a single bin, i.e., 2-DIMENSIONAL KNAPSACK, admits a PTAS due to Frieze and Clarke [9]. As shown in [14], an EPTAS for the problem is unlikely to exist. The first PTAS for MULTIPLE KNAPSACK was presented by Chekuri and Khanna [6] who

[1] We say that $(a_1, a_2) \leq (b_1, b_2)$ if $a_1 \leq b_1$ and $a_2 \leq b_2$.
[2] This is also known as virtual machine *instantiation* [5].
[3] We give the proof of Theorem 1 in Section 4.

also showed the problem is strongly NP-hard. The PTAS was later improved to an EPTAS by Jansen [10, 11]. For comprehensive surveys of known results on knapsack problems, see, e.g., [12, 4].

Both Multiple Knapsack and 2VMK are special cases of the Separable Assignment Problem (SAP) studied in [8]. The input for SAP is a set of items I and m bins. Each item $i \in I$ has a profit $p_{i,j}$ if assigned to bin $j \in [m]$. Each bin $j \in [m]$ is associated with a collection of feasible assignments $\mathcal{F}_j \subseteq 2^I$. The feasible assignments are hereditary, that is, if $S \in \mathcal{F}_j$ and $T \subseteq S$ then $T \in \mathcal{F}_j$ for all $T \subseteq S$. The feasible assignments are given implicitly via a β-optimization oracle which, given a value function $v : I \to \mathbb{R}_{\geq 0}$ and $j \in [m]$, finds a β-approximate solution for $\max_{S\in\mathcal{F}_j} \sum_{i\in S} v(i)$. A solution for the SAP instance is a tuple of disjoint sets $S_1, \ldots, S_m \subseteq I$ such that $S_j \in \mathcal{F}_j$ for all $j \in [m]$. The objective is to find a solution $S_1, \ldots, S_m$ of maximum profit $\sum_{j=1}^{m} \sum_{i\in S_j} p_{i,j}$. A $(1 - e^{-1}) \cdot \beta$-approximation for SAP was given in [8].

Observe that 2VMK can be cast as SAP by setting $p_{i,j} = p(i)$ for every $(i, j) \in I \times [m]$, and $\mathcal{F}_j = \{S \subseteq I \mid w(S) \leq (1, 1)\}$ for every $j \in [m]$. That is, the profit of item i is $p(i)$ regardless of the bin to which it is assigned, and the feasible assignments of all bins are simply all subsets of items which fit into a bin. For every fixed $\varepsilon > 0$, a $(1 - \varepsilon)$-optimization oracle for the bins can be implemented in polynomial time using the PTAS of [9] for 2-dimensional knapsack. Hence, a $(1 - e^{-1} - \varepsilon)$-approximation for 2VMK follows from the result of [8].

Parameterized algorithms for 2VMK were proposed in [15, 1]. We are not aware of earlier works which directly study 2VMK from approximation algorithms viewpoint.

1.2 Technical Overview

Our algorithm combines the approximation algorithm of Fleischer et al. [8] with a simple reduction of 2VMK to (1-dimensional) Multiple Knapsack.

The algorithm of [8] is based on randomized rounding of a *configuration-LP* solution. Given an instance $\mathcal{I} = (I, w, p, m)$ of 2VMK, let $w_j(i)$ denote the weight of item i in the jth coordinate, for all $i \in I$ and $j \in \{1, 2\}$. Also, given $f : I \to \mathbb{R}^d$ we use the notation $f(S) = \sum_{i\in S} f(i)$ for all $S \subseteq I$. A *configuration* of the instance $\mathcal{I}$ is $C \subseteq I$ such that $w(C) \leq (1, 1)$. Let $\mathcal{C}(\mathcal{I})$ be the set of all configurations of the 2VMK instance $\mathcal{I}$. For any $i \in I$, let $\mathcal{C}(\mathcal{I}, i)$ be the set of configurations containing item i. We often omit $\mathcal{I}$ and use $\mathcal{C}$ and $\mathcal{C}(i)$ when the instance $\mathcal{I}$ is known by context. Observe that a solution for the instance $\mathcal{I}$ is a tuple of m configurations.

Given a 2VMK instance $\mathcal{I} = (I, w, p, m)$, let $\mathbf{x}_C \in \{0, 1\}$ be an indicator for the selection of configuration $C \in \mathcal{C} = \mathcal{C}(\mathcal{I})$ for the solution. In the configuration-LP relaxation of the problem, we have $\mathbf{x}_C \geq 0\ \forall C \in \mathcal{C}$. Our algorithm initially solves the following.

$$
\begin{aligned}
&\text{(C-LP)} \quad && \max && \sum_{C\in\mathcal{C}} \sum_{i\in C} \mathbf{x}_C \cdot p(i) && \\
& && \text{subject to :} && \sum_{C\in\mathcal{C}} \mathbf{x}_C \leq m && && (1)\\
& && && \sum_{C\in\mathcal{C}(i)} \mathbf{x}_C \leq 1 && \forall i \in I && (2)\\
& && && \mathbf{x}_C \geq 0 && \forall C \in \mathcal{C} &&
\end{aligned}
$$

The next lemma follows from a result of [8].

▶ **Lemma 3.** *There is a PTAS for* C-LP*.*

Let $\mathbf{x}$ be a solution for C-LP. We say that a random configuration $R \in \mathcal{C}$ is *distributed by* $\mathbf{x}$ if $\Pr[R = C] = \frac{\mathbf{x}_C}{m}$.[4]

To obtain a $\left((1-e^{-1}) \cdot (1-\varepsilon)\right)$-approximation for 2VMK, the algorithm of [8] finds a $(1-\varepsilon)$-approximate solution $\mathbf{x}$ for C-LP, and then independently samples m configuration $R_1, \ldots, R_m \in \mathcal{C}$, where each of the configurations R_j is distributed by $\mathbf{x}$. The returned solution is the sampled configurations $R_1, \ldots, R_m$. For simplicity of this informal overview, assume that $\sum_{C \in \mathcal{C}} \sum_{i \in C} \mathbf{x}_C \cdot p(i) \approx \mathrm{OPT}(\mathcal{I})$.

It can be shown that

$$\mathbb{E}\left[p(R_1 \cup \ldots \cup R_j)\right] \approx \left(1 - e^{-\frac{j}{m}}\right) \cdot \sum_{C \in \mathcal{C}} \sum_{i \in C} \mathbf{x}_C \cdot p(i) \approx \left(1 - e^{-\frac{j}{m}}\right) \cdot \mathrm{OPT}(\mathcal{I}) \tag{3}$$

for all $j \in [m]$, and in particular $\mathbb{E}\left[p(R_1 \cup \ldots \cup R_m)\right] \approx \left(1 - e^{-1}\right) \cdot \mathrm{OPT}(\mathcal{I})$. Define

$$q_j = \mathbb{E}\left[p(R_1 \cup \ldots \cup R_j) - p(R_1 \cup \ldots \cup R_{j-1})\right]$$

to be the marginal expected profit of the jth sampled configuration. By (3) we have

$$q_j \approx \left(\left(1 - e^{-\frac{j}{m}}\right) - \left(1 - e^{-\frac{j-1}{m}}\right)\right) \cdot \mathrm{OPT}(\mathcal{I}) \approx \frac{1}{m} \cdot e^{-\frac{j}{m}} \cdot \mathrm{OPT}(\mathcal{I}), \tag{4}$$

where the last estimation follows from $e^{-\frac{j-1}{m}} \approx e^{-\frac{j}{m}} + \frac{1}{m} \cdot e^{-\frac{j}{m}}$ by a Taylor expansion of e^x at $x = -\frac{j}{m}$. Equation (4) implies that the marginal profit from each configuration decreases as the sampling process proceeds. The first sampled configuration has a marginal profit of $\approx \frac{1}{m} \cdot \mathrm{OPT}(\mathcal{I})$, while the marginal profit of the mth configuration is $\approx e^{-1} \cdot \frac{1}{m} \cdot \mathrm{OPT}(\mathcal{I}) \approx 0.367 \cdot \frac{1}{m} \cdot \mathrm{OPT}(\mathcal{I})$.

We note that a simple reduction to Multiple Knapsack can be used to derive a $\left(\frac{1}{2} - \varepsilon\right)$-approximation for 2VMK. An instance $\mathcal{J}$ of uniform multiple knapsack (MK) is a tuple $\mathcal{J} = (I, w, p, m)$, where I is a set of items, $w : I \to [0,1]$ is a weight function, $p : I \to \mathbb{R}_{\geq 0}$ is a profit function, and m is the number of (unit size) bins. A *configuration* of $\mathcal{J}$ is a subset of items $C \subseteq I$ satisfying $w(C) \leq 1$. We use $\mathcal{C}(\mathcal{J})$ to denote the set of all configurations of $\mathcal{J}$, and sometimes omit $\mathcal{J}$ from this notation if it is known by context. A feasible solution for $\mathcal{J}$ is a tuple of m configurations $C_1, \ldots, C_m \in \mathcal{C}(\mathcal{J})$. The objective is to find a solution $C_1, \ldots C_m$ for which the total profit, given by $p(\bigcup_{t \in [m]} C_t)$, is maximized. We note that uniform MK can be viewed as uniform 1-dimensional vector multiple knapsack.

Let $\mathcal{I} = (I, w, p, m)$ be a 2VMK instance, and define an MK instance $\mathcal{J} = (I, w', p, m)$ where $w'(i) = \max\{w_1(i), w_2(i)\}$ for every $i \in I$. It can be easily shown that $\mathcal{C}(\mathcal{J}) \subseteq \mathcal{C}(\mathcal{I})$, i.e., a configuration of the MK instance is a configuration of the 2VMK instance. Furthermore, every $C \in \mathcal{C}(\mathcal{I})$ can be partitioned into $C_1, C_2 \in \mathcal{C}(\mathcal{J})$ (possibly with $C_1 = \emptyset$). That is, every configuration of the 2VMK instance can be split into two configurations of the MK instance. Therefore $\mathrm{OPT}(\mathcal{J}) \geq \frac{1}{2} \cdot \mathrm{OPT}(\mathcal{I})$, as we can take the optimal solution of the 2VMK instance $\mathcal{I}$, convert its configurations to $2m$ configurations of the MK instance $\mathcal{J}$, and select the m most profitable ones. As MK admits a PTAS [6], we can find a $(1-\varepsilon)$-approximate solution for $\mathcal{J}$ in polynomial time. This leads to the following algorithm, to which we refer as the *reduction algorithm*. Given the instance $\mathcal{I}$, return a $(1-\varepsilon)$-approximate solution for $\mathcal{J}$. The above arguments can be used to show that the reduction algorithm is a $\left(\frac{1}{2} - \varepsilon\right)$-approximation algorithm for 2VMK.

[4] W.l.o.g we assume that $\|x\| = \sum_{C \in \mathcal{C}} \mathbf{x}_C = m$.

While this simple $\left(\frac{1}{2}-\varepsilon\right)$-approximation is inferior to the $\left(1-e^{-1}-\varepsilon\right)$-approximation of Fleischer et al. [8], it is still useful for obtaining a better approximation for 2VMK, due to the following property. Let $T \subseteq I$ be a random subset of the items such that $\Pr(i \in T) \leq \alpha$ for all $i \in I$, and T satisfies some concentration bounds. Then, if the reduction algorithm is executed with the 2VMK instance $(I \setminus T, p, w, (1-\alpha)m)$, it returns with high probability a solution of profit at least $\frac{1}{2} \cdot (1-\alpha) \cdot \mathrm{OPT}(\mathcal{I})$. In other words, if every item is removed from the instance with probability at most α, then the algorithm can find a packing into $(1-\alpha) \cdot m$ bins that yields $\frac{1}{2} \cdot (1-\alpha)$ of the original profit $\mathrm{OPT}(\mathcal{I})$. This property resembles the notion of *subset oblivious* algorithms used by [2] as part of their Round&Approx framework for set covering problems. Indeed, we use a subset oblivious algorithm of [2] in our proof of this property.

This suggests the following hybrid algorithm. First, solve C-LP and sample $\ell \approx \alpha m$ configuration $R_1, \ldots, R_\ell$ distributed by the C-LP solution $\mathbf{x}$. Subsequently, define $T = \bigcup_{j=1}^{\ell} R_j$, and use the reduction algorithm to find a solution S for the instance $(I \setminus T, w, p, (1-\alpha)m)$. The algorithm returns $R_1, \ldots, R_\ell$ together with the $m - \ell = (1-\alpha)m$ configurations of S. It can be shown that $\Pr(i \in T) \leq \alpha$ for all T, and that T satisfies the required concentration bounds; therefore, the expected profit of S is $\approx \frac{1}{2} \cdot (1-\alpha) \cdot \mathrm{OPT}(\mathcal{I})$. Since S uses $(1-\alpha)m$ configurations, this can be interpreted as a residual profit of $\frac{1}{2 \cdot m} \cdot \mathrm{OPT}(\mathcal{I})$ per configuration. This property suggests how to select α. We set the value of α such that the marginal profit of the (αm)-th sampled configuration is equal to $\frac{1}{2m} \cdot \mathrm{OPT}(\mathcal{I})$, the marginal profit of a configuration in S. By (4), this is realized for $\alpha = \ln 2$.

We note that this hybrid approach is similar to the Round&Approx framework of Bansal et al. [2], which solves set covering problems by sampling random configurations based on a solution for a configuration-LP, followed by a subset oblivious algorithm which completes the solution. Our algorithm can be viewed as an adaptation of the framework of [2] to knapsack variants. To the best of our knowledge, this is the first application of ideas from [2] to maximization problems.

Our proofs rely on a dimension-free concentration bound for self-bounding functions due to Boucheron et al. [3] (see Section 3.1 for details). While our approach is conceptually similar to the Round&Approx of [2], which uses McDiarmid's bound [16] to show concentration, it appears that McDiarmid's bound does not suffice to guarantee concentration of the profits. The bound of Boucheron et al. was previously used by Vondrák to show concentration bounds for submodular functions [20]. We are not aware of other applications of this bound in the context of combinatorial optimization.

1.3 Organization

In Section 2 we give some definitions and preliminary results. Section 3 presents our approximation algorithm for 2VMK, and in Section 4 we give a proof of APX-hardness (stated in Theorem 1). We conclude in Section 5 with a summary and some directions for future work. Due to space constraints, some of the proofs are given in the full version of the paper [7].

2 Preliminaries

For any function $f : I \to \mathbb{R}^d$ where $d \in \{1, 2\}$ we use the notation $f(A) = \sum_{i \in A} f(i)$. We note that in both MK and 2VMK an item selected for the solution may appear in more than one configuration; however, the profit of each selected item is counted exactly once.

Let $\mathcal{I}$ be an instance of either 2VMK or MK. We say that a subset of items $S \subseteq I$ *can be packed into q bins of $\mathcal{I}$*, for some $1 \leq q \leq m$, if there are configurations $C_1, \ldots, C_q \in \mathcal{C}(\mathcal{I})$ such that $\cup_{t=1}^{q} C_t = S$.

Finally, given $n \in \mathbb{N}$ and an arbitrary set $\mathcal{X}$, define $\mathcal{X}^n$ to be the set of all vectors of dimension n over $\mathcal{X}$; that is, $\mathcal{X}^n = \{(x_1, \ldots, x_n) \mid \forall t \in \{1, \ldots, n\} : \; x_t \in \mathcal{X}\}$.

2.1 Associated Instances

Our approximation algorithm for 2VMK uses as a subroutine an approximation algorithm for MK. To this end, we define for a given 2VMK instance an associated MK instance. Formally, given a 2VMK instance $\mathcal{I} = (I, w, p, m)$, we define the *$k$-associated MK instance* $\mathcal{I}' = (I', w', p', m')$ as follows. The set of items is $I' = I$. The item weights are given by $w'(i) = \max\{w_1(i), w_2(i)\}$, for $1 \leq i \leq n$; the profit of item i is $p'(i) = p(i)$, and the number of (unit size) bins is $m' = k \cdot m$. The next result will be useful in analyzing our algorithm for 2VMK.

▶ **Lemma 4.** *Let $\mathcal{I} = (I, w, p, m)$ be a 2VMK instance, and $\mathcal{I}' = (I', w', p', m')$ its k-associated instance. Then, any set $S \subseteq I$ that can be packed in q bins of $\mathcal{I}$, can be packed in $2 \cdot q$ bins of $\mathcal{I}'$.*

2.2 Restriction to ε-Nice Instances

The correctness proofs for our approach require the number of bins m to be large, and the maximum profit for a single configuration to be relatively small. These properties are essential for the concentration bounds that we use to ensure success with high probability. We show that for instances in which m is small (i.e., bounded by some constant), a simple reduction to d-dimensional knapsack with a matroid constraint yields a PTAS. This allows us to focus on instances with a large number of bins. Furthermore, for such instances we use a simple greedy pre-processing to ensure bounded maximal profit for a single configuration. Thus, we restrict our attention to the following subclass of *ε-nice* instances. Let $\exp^{[k]}(x) = \exp(\exp^{[k-1]}(x))$, for any integer $k \geq 2$, and $\exp^{[1]}(x) = \exp(x)$.

▶ **Definition 5.** *Given $\varepsilon \in (0, 0.01)$, an instance $\mathcal{I} = (I, w, p, m)$ is* ε-nice *if $m \geq \exp^{[3]}(\varepsilon^{-30})$ and $p(C) \leq \varepsilon^{20} \cdot \mathrm{OPT}(\mathcal{I})$ for every $C \in \mathcal{C}$.*[5]

We show that efficient approximation for 2VMK on ε-nice instances yields almost the same approximation ratio for general instances.

▶ **Lemma 6.** *For any $\varepsilon \in (0, 0.01)$ and $\beta \in (0, 1 - \varepsilon)$, if there is polynomial-time β-approximation algorithm for* 2VMK *on ε-nice instances, then there is a polynomial-time $(1 - \varepsilon) \cdot \beta$-approximation algorithm for* 2VMK.

2.3 Two-dimensional Vector Bin Packing

An instance $\mathcal{I}$ of the 2-Dimensional Vector Bin Packing (2VBP) problem is a pair (I, w), where I is a set of n items and $w : I \to [0, 1]^2$ is a two-dimensional weight function. A *solution* for the instance (I, w) is a collection of subsets of items $S_1, \ldots, S_m \subseteq I$ such that $w(S_b) = \sum_{i \in S_b} w(i) \leq (1, 1)$ for all $b = 1, \ldots, m$, and $\bigcup_{b=1}^{m} S_b = I$. The *size* of the solution is m. Our objective is to find a solution of minimum size.

[5] We did not attempt to optimize the constants.

An *asymptotic polynomial-time approximation scheme* (APTAS) is an infinite family $\{A_\varepsilon\}$ of asymptotic $(1-\varepsilon)$-approximation algorithms, one for each $\varepsilon > 0$. Ray [17] showed that 2VBP does not admit an asymptotic approximation ratio better than $\frac{600}{599}$, assuming $P \neq NP$; thus, 2VBP does not admit an APTAS.

3 Approximation Algorithm for ε-Nice Instances

In this section we present an algorithm for ε-nice 2VMK instances. Our algorithm proceeds by initially obtaining an approximate solution $\mathbf{x}$ for C-LP (as given in Section 1.2), and then forming a partial solution by sampling $1 \leq \ell \leq m$ configurations. The remaining $(m-\ell)$ configurations are derived by solving the associated MK instance for the remaining (unassigned) items. The pseudocode of our algorithm is given in Algorithm 1.

Algorithm 1 Approximation Algorithm for ε-nice instances.

configuration: $\varepsilon \in (0, 0.01)$.
input : An ε-*nice* instance $\mathcal{I} = (I, w, p, m)$ of 2VMK.
output : A solution for the instance $\mathcal{I}$.
1: Find a $(1-\varepsilon)$-approximate solution $\mathbf{x}$ for C-LP; let x^* be its value.
2: **for** $t = 1$ *to* $\ell = \lceil m \cdot \ln 2 \rceil$ **do**
   Sample a random configuration R_t distributed by $\mathbf{x}$
   end
3: $S \leftarrow I \setminus (\cup_{t \in \{1,\ldots,l\}} R_t)$
4: Let $\mathcal{I}'$ be the 1-associated MK instance of the 2VMK instance $(S, w, p, m-\ell)$.
5: Find a $(1-\varepsilon)$-approximate solution for the MK instance $\mathcal{I}'$; denote the solution by $R_{\ell+1}, \ldots, R_m$.
6: Return $(R_1, \ldots, R_m)$.

Note that, by Lemma 3, Step 1 of Algorithm 1 can be implemented in polynomial time, for any fixed $\varepsilon > 0$. Let $\varepsilon \in (0, 0.01)$. and $\mathcal{I}$ be an ε-nice 2-VMK instance. Consider the execution of Algorithm 1 configured by ε with $\mathcal{I}$ as its input. Let OPT be the set of items selected by an optimal solution for $\mathcal{I}$, and $T = \bigcup_{t \in \{1,\ldots,\ell\}} R_t$ the items selected in Step 2 in Algorithm 1. We use the next lemmas in the analysis of the algorithm. Lemma 7 lower bounds the expected profit of $R_1, \ldots, R_\ell$, the configurations sampled in Step 2 of Algorithm 1. Lemma 8 gives a lower bound on the profit of the MK solution $R_{\ell+1}, \ldots, R_m$ found in Step 5 of Algorithm 1. Lemma 9 lower bounds the profit of the solution returned by the algorithm using the bounds in Lemmas 7 and 8, whose proofs are given in Sections 3.2 and 3.3.

▶ **Lemma 7.** $\Pr\left[p(T) \leq (1 - e^{-\alpha} - 2 \cdot \varepsilon) \cdot p(\mathrm{OPT})\right] \leq \exp\left(-\varepsilon^{-7}\right)$.

▶ **Lemma 8.** $\Pr\left[p(\bigcup_{t=\ell+1}^{m} R_t) \leq \left(\frac{1-\alpha}{2} - 3 \cdot \varepsilon\right) \cdot p(\mathrm{OPT})\right] \leq \frac{1}{4}$.

▶ **Lemma 9.** *Algorithm 1 returns a solution of profit at least* $\left(1 - \frac{\ln 2}{2} - 5 \cdot \varepsilon\right) \cdot p(\mathrm{OPT})$ *with probability at least* $\frac{1}{2}$.

For an event A, let $\bar{A}$ denote the complementary event.

Proof of Lemma 9. Let A be the event "$p\left(T\right) > (1 - e^{-\alpha} - 2 \cdot \varepsilon) \cdot p(\mathrm{OPT})$", and B the event "$p\left(\bigcup_{t=\ell+1}^{m} R_t\right) > \left(\frac{1-\alpha}{2} - 3\varepsilon\right) \cdot p(\mathrm{OPT})$". If both A and B occur then Algorithm 1 returns a solution of profit at least

$$p(T) + p\left(\bigcup_{t=\ell+1}^{m} R_t\right) \geq \left(1 - e^{-\alpha} - 2\varepsilon + \frac{1-\alpha}{2} - 3\varepsilon\right) p(\mathrm{OPT}) \quad = \left(1 - \frac{\ln 2}{2} - 5\varepsilon\right) p(\mathrm{OPT}).$$

The inequality holds since both A and B occur. The equality holds since $\alpha = \ln 2$. The probability that A and B occur is given by

$$\Pr\left[A \cap B\right] = 1 - \Pr\left[\bar{A} \cup \bar{B}\right] \quad \geq 1 - \left(\Pr\left[\bar{A}\right] + \Pr\left[\bar{B}\right]\right) \geq 1 - \exp\left(-\varepsilon^{-7}\right) - \frac{1}{4} \geq \frac{1}{2}$$

The first inequality follows from the union bound. The second inequality follows from Lemmas 7 and 8, and since $\varepsilon < 0.01$. ◀

3.1 Self-Bounding Functions

Lemmas 7 and 8 we use a concentration bound for *self-bounding functions.*

▶ **Definition 10.** *A non-negative function $f : \mathcal{X}^n \to \mathbb{R}_{\geq 0}$ is called self-bounding if there exist n functions $f_1, \ldots, f_n : \mathcal{X}^{n-1} \to \mathbb{R}$ such that for all $x = (x_1, \ldots, x_n) \in \mathcal{X}^n$,*

$$0 \leq f(x) - f_t(x^{(t)}) \leq 1, \qquad \text{and}$$

$$\sum_{t=1}^{n} \left(f(x) - f_t(x^{(t)})\right) \leq f(x),$$

where $x^{(t)} = (x_1, \ldots, x_{t-1}, x_{t+1}, \ldots, x_n) \in \mathcal{X}^{n-1}$ is obtained by dropping the t-th component of x.

The next result is shown in [3].

▶ **Lemma 11.** *Let $f : \mathcal{X}^n \to \mathbb{R}_{\geq 0}$ be a self-bounding function and let $X_1, \ldots, X_n \in \mathcal{X}$ be independent random variables. Define $Z = f(X_1, \ldots, X_n)$. Then the following holds:*

1. $\Pr[Z \geq \mathbb{E}[Z] + t] \leq \exp\left(-\frac{t^2}{2 \cdot \mathbb{E}[Z] + \frac{t}{3}}\right)$*, for every $t \geq 0$.*
2. $\Pr[Z \leq \mathbb{E}[Z] - t] \leq \exp\left(-\frac{t^2}{2 \cdot \mathbb{E}[Z]}\right)$*, for every $0 < t < \mathbb{E}[Z]$.*

We use the following construction of self-bounding functions several times in the paper. Recall that $\mathcal{C} = \mathcal{C}(\mathcal{I})$.

▶ **Lemma 12.** *Let $\mathcal{I} = (I, w, p, m)$ be a dVMK instance, and $h : I \to \mathbb{R}_{\geq 0}$. Define $f : \mathcal{C}^\ell \to \mathbb{R}$ by $f(C_1, \ldots, C_\ell) = \frac{h(\bigcup_{i \in [\ell]} C_i)}{\eta}$, where $\eta \geq \max_{C \in \mathcal{C}} h(C)$. Then f is a self-bounding function.*

3.2 Profit of the Sampled Configurations

In this section we prove Lemma 7; namely, we show that with high probability the profit $p(\cup_{t \in \{1,\ldots,\ell\}} R_t)$ is sufficiently large. We first prove the next lemma.

▶ **Lemma 13.** $\mathbb{E}[p(T)] \geq (1 - e^{-\alpha} - \varepsilon) \cdot p(\mathrm{OPT})$.

Proof. Let $\mathcal{C}(i) = \mathcal{C}(\mathcal{I}, i)$ be the set of configurations containing item $i \in \mathcal{I}$, then for every $i \in I$, the probability that i is not contained in the sampled configurations is

$$\Pr[i \notin T] = \Pr[i \notin \cup_{t \in \{1,\ldots,\ell\}} R_t] = \prod_{t \in \{1,\ldots,\ell\}} \Pr[i \notin R_t] = \left(1 - \sum_{C \in \mathcal{C}(i)} \frac{x_C}{m}\right)^{\ell},$$

where third equality holds since $\Pr[i \in R_t] = \sum_{C \in \mathcal{C}(i)} \frac{x_C}{m}$, for all $t \in \{1, \ldots, \ell\}$. Hence,

$$\Pr[i \notin T] = \left(1 - \sum_{C \in \mathcal{C}(i)} \frac{x_C}{m}\right)^{\ell} \leq \left(1 - \sum_{C \in \mathcal{C}(i)} \frac{x_C}{m}\right)^{m \cdot \alpha \cdot \frac{\sum_{C \in \mathcal{C}(i)} x_C}{\sum_{C \in \mathcal{C}(i)} x_C}} \leq \exp\left(-\alpha \cdot \sum_{C \in \mathcal{C}(i)} x_C\right).$$

The first inequality holds since $\ell \geq \alpha \cdot m$. The second inequality holds by $(1-\frac{1}{x})^x \leq e^{-1}$ for $x \geq 1$. Thus, we have

$$\Pr[i \in T] = 1 - \Pr[i \notin T] \quad \geq \left(1 - \exp\left(-\alpha \cdot \sum_{C \in \mathcal{C}(i)} x_C\right)\right) \geq \sum_{C \in \mathcal{C}(i)} x_C \cdot (1 - e^{-\alpha})$$

For the second inequality, we used $1 - e^{-x \cdot \alpha} \geq x \cdot (1 - e^{-\alpha})$ for $x, \alpha \in [0,1]$. Therefore,

$$\begin{aligned}
\mathbb{E}[p(T)] &= \sum_{i \in I} p(i) \cdot \Pr[i \in T] \\
&\geq \sum_{i \in I} p(i) \cdot \sum_{C \in \mathcal{C}(i)} x_C \cdot (1 - e^{-\alpha}) \\
&= (1 - e^{-\alpha}) \cdot \sum_{C \in \mathcal{C}} \sum_{i \in C} x_C \cdot p(i) \\
&= (1 - e^{-\alpha}) \cdot x^* \\
&\geq (1 - e^{-\alpha} - \varepsilon) \cdot p(\mathrm{OPT}).
\end{aligned}$$

The third equality follows from our definition of x^* as the value of the solution $\mathbf{x}$ found in Step 1 of Algorithm 1. The second inequality holds since $x^* \geq (1-\varepsilon) \cdot p(\mathrm{OPT})$. ◀

Proof of Lemma 7. Define $f : C^\ell \to \mathbb{R}$ by $f(C_1, \ldots, C_\ell) = \frac{p(\bigcup_{t \in \{1,\ldots,\ell\}} C_t)}{\varepsilon^{10} \cdot p(\mathrm{OPT})}$ As the instance $\mathcal{I}$ is ε-nice, we have that $\varepsilon^{10} \cdot p(\mathrm{OPT}) \geq \max_{C \in \mathcal{C}} p(C)$. By Lemma 12, f is a self-bounding function. Hence, by Lemma 13 we have,

$$\begin{aligned}
&\Pr\left[p(T) \leq (1 - e^{-\alpha} - 2 \cdot \varepsilon) \cdot p(\mathrm{OPT})\right] \\
&\quad \leq \Pr\left[\frac{p(T)}{\epsilon^{10} \cdot p(OPT)} \leq \frac{\mathbb{E}[p(T)]}{\epsilon^{10} \cdot p(\mathrm{OPT})} - \frac{\varepsilon \cdot p(\mathrm{OPT})}{\epsilon^{10} \cdot p(\mathrm{OPT})}\right] \\
&\quad = \Pr\left[f(R_1, \ldots, R_\ell) \leq \mathbb{E}[f(R_1, \ldots, R_\ell)] - \varepsilon^{-9}\right] \\
&\quad \leq \exp\left(-\frac{\varepsilon^{-18}}{2 \cdot \mathbb{E}[f(R_1, \ldots, R_\ell)]}\right) \\
&\quad \leq \exp\left(-\frac{\varepsilon^{-18}}{2 \cdot \frac{p(\mathrm{OPT})}{\epsilon^{10} \cdot p(\mathrm{OPT})}}\right) \leq \exp\left(-\varepsilon^{-7}\right).
\end{aligned}$$

The first equality holds by the definition of f. The second inequality follows from Lemma 11, by taking $t = \varepsilon^{-9}$. The third inequality holds since $f(R_1, \ldots, R_\ell) \leq \frac{p(\mathrm{OPT})}{\varepsilon^{10} \cdot p(\mathrm{OPT})}$, as $R_1, \ldots, R_\ell$ along with additional $m - \ell$ empty configurations is a solution for $\mathcal{I}$. The fourth inequality holds since $2 \cdot \varepsilon \leq 1$. ◀

3.3 The Solution for the Residual Items

In this section we prove Lemma 8. Specifically, we show that the profit of the solution for the MK instance constructed in Step 4 of Algorithm 1 is sufficiently high. Since we obtain a $(1-\varepsilon)$-approximate solution for the MK instance $\mathcal{I}'$, we only need to derive a lower bound for $\mathrm{OPT}(\mathcal{I}')$. To this end, we show that there exists a set $Q \subseteq \mathrm{OPT}$, such that the set $Q \setminus T$ has sufficiently high profit $p(Q \setminus T)$, and $Q \setminus T$ can be almost entirely packed in twice the number of remaining bins. We choose among these bins the most profitable ones to obtain the lower bound. In our analysis, we use the notion of subset-obliviousness, introduced in [2]. The following is a simplified version of a definition given in [2] w.r.t. the BIN PACKING (BP) problem. Let `BP-OPT`(I, w) denote the size of an optimal solution for a BP instance $\mathcal{I} = (I, w)$.

▶ **Definition 14.** *Let $\rho > 1$. We say that* Bin Packing *is ρ-subset oblivious if, for any fixed $\varepsilon > 0$, there exist k, ψ, δ (possibly depending on ε) such that, for any BP instance $\mathcal{I} = (I, w)$, there exist functions $g_1, \ldots, g_k : 2^I \to \mathbb{R}_{\geq 0}$ which satisfy the following.*

(i) $g_t(C) \leq \psi$ *for any* $C \in \mathcal{C}(\mathcal{I})$ *and* $t \in \{1, \ldots, k\}$*;*

(ii) $\texttt{BP-OPT}(I, w) \geq \max_{t \in \{1,\ldots,k\}} g_t(I)$*;*

(iii) $\texttt{BP-OPT}(S, w) \leq \rho \cdot \max_{t \in \{1,\ldots,k\}} g_t(S) + \varepsilon \cdot \texttt{BP-OPT}(I, w) + \delta$*, for all* $S \subseteq I$*.*

We refer to the values k, ψ and δ as the (ρ, ε)-subset oblivious parameters of Bin Packing*, and the functions $g_1, \ldots, g_k$ as the (ρ, ε)-subset oblivious functions of $\mathcal{I}$.*

The next lemma follows from a result of [2].

▶ **Lemma 15.** *For any fixed $\varepsilon > 0$,* Bin Packing *is $(1+\varepsilon)$-subset oblivious, and the $(1+\varepsilon, \varepsilon)$ parameters k, δ, ψ satisfy $k \leq \exp^{[3]}(\varepsilon^{-1})$, $\delta \leq \frac{4}{\varepsilon^4}$, and $\psi \leq 1$.*

Let $\mathcal{J} = (\mathrm{OPT}, w', p, 2 \cdot m)$ be the 2-associated MK instance of $\mathcal{I}_{\mathrm{OPT}} = (\mathrm{OPT}, w, p, m)$. By Lemma 15, Bin Packing is $(1+\varepsilon^2)$-subset oblivious. Thus, there exist k, ψ, δ which are $(1+\varepsilon^2, \varepsilon^2)$ subset oblivious parameters of Bin Packing. Let $g_1, \ldots, g_k$ be the $(1+\varepsilon^2, \varepsilon^2)$ subset-oblivious functions of the Bin Packing instance (OPT, w'). By Lemma 15, the values k, ψ and δ satisfy, $k \leq \exp^{[3]}(\varepsilon^{-2})$, $\delta \leq \frac{4}{\varepsilon^8}$, $\psi \leq 1$. Define $\alpha' = \frac{\lceil \alpha \cdot m \rceil}{m}$, then as $\alpha' \cdot m - \alpha \cdot m \leq 1$, we have that $\alpha' - \alpha \leq \frac{1}{m} \leq \varepsilon$.

▶ **Lemma 16.** *There exists $Q \subseteq \mathrm{OPT}$ which satisfies the following.*

1. $\Pr\left[g_t(Q \setminus T) \geq (1-\alpha') \cdot g_t(\mathrm{OPT}) + k \cdot \psi + \varepsilon^{10} \cdot m\right] \leq \exp\left(-\frac{\varepsilon^{21} \cdot m}{\psi^2}\right)$*, for all* $t \in \{1, \ldots, k\}$*.*
2. $\Pr\left[p(Q \setminus T) \leq (1-\alpha'-\varepsilon) \cdot p(\mathrm{OPT})\right] \leq \exp\left(-\varepsilon^{-7}\right)$*.*

Proof. To show the existence of the set Q satisfying the properties in the lemma, consider first the following optimization problem. Given an optimal solution OPT for a 2VMK instance $\mathcal{I}$, find a subset of items $Q \in \mathrm{OPT}$ for which $\mathbb{E}\left[p(Q \setminus T)\right]$ is maximized, under the constraint that $\mathbb{E}\left[g_t(Q \setminus T)\right] \leq (1-\alpha')g_t(\mathrm{OPT})$ for all $t \in [k]$. Let $y_i \in \{0, 1\}$ be an indicator for the inclusion of item $i \in \mathrm{OPT}$ in Q. We can formulate an integer program for the above optimization problem. In the following LP relaxation we have $0 \leq y_i \leq 1$, $\forall\, i \in \mathrm{OPT}$.

$$
\begin{array}{lll}
(Q\text{-LP}) \quad \max & \displaystyle\sum_{i \in \mathrm{OPT}} y_i \cdot p(i) \cdot \Pr[i \notin T] & \\
\text{s.t.} & \displaystyle\sum_{i \in \mathrm{OPT}} y_i \cdot \Pr[i \notin T] \cdot g_t(i) \leq (1-\alpha')g_t(\mathrm{OPT}) & \forall t \in \{1, \ldots, k\} \qquad (5) \\
& 0 \leq y_i \leq 1 & \forall i \in \mathrm{OPT}
\end{array}
$$

Let $\mathbf{y}^*$ be a basic optimal solution for Q-LP. We define

$$Q = \{i \in Q \mid y_i^* > 0\} \tag{6}$$

to be the set of all items with positive entries in $\mathbf{y}^*$. We show that the set Q defined in (6) satisfies $\mathbb{E}\left[p(Q \setminus T)\right] \geq (1-\alpha')p(\mathrm{OPT})$. To this end, we prove the next claim.

▷ Claim 17. $\sum_{i \in \mathrm{OPT}} y_i^* \cdot p(i) \cdot \Pr[i \notin T] \geq (1-\alpha') \cdot p(\mathrm{OPT})$.

Proof. For every $i \in \mathrm{OPT}$, the following holds:

$$
\begin{aligned}
\Pr[i \in T] &= \Pr\left[\exists t \in \{1, \ldots, \ell\}, i \in R_t\right] \\
&\leq \sum_{t \in \{1,\ldots,\ell\}} \Pr\left[i \in R_t\right] \\
&= \sum_{t \in \{1,\ldots,\ell\}} \sum_{C \in \mathcal{C}(i)} \frac{x_C}{m} \leq \sum_{t \in \{1,\ldots,\ell\}} \frac{1}{m} = \frac{\ell}{m}.
\end{aligned}
$$

Consider the vector $\mathbf{y}' = (y'_1, \ldots, y'_{|\mathrm{OPT}|})$ where $y'_i = \dfrac{1-\alpha'}{\Pr[i \notin T]}$ for every $i \in \mathrm{OPT}$. Then $\mathbf{y}'$ is a feasible solution for Q-LP since the following holds:

1. For every $i \in \mathrm{OPT}$, $y'_i = \frac{1-\alpha'}{\Pr[i \notin T]}$ satisfies the following,
$$0 \leq \frac{1-\alpha'}{\Pr[i \notin T]} = \frac{1-\alpha'}{1-\Pr[i \in T]} \leq \frac{1-\alpha'}{1-\frac{\ell}{m}} = \frac{1-\alpha'}{1-\frac{\alpha' \cdot m}{m}} = 1.$$
Therefore $0 \leq y'_i \leq 1$, for all $i \in \mathrm{OPT}$.
2. For every $t \in \{1, \ldots, k\}$, the following holds:
$$\sum_{i \in \mathrm{OPT}} y'_i \cdot \Pr[i \notin T] \cdot g_t(i) = (1-\alpha') \cdot \sum_{i \in \mathrm{OPT}} g_t(i) = (1-\alpha') \cdot g_t(\mathrm{OPT}).$$

The objective value $\sum_{i \in \mathrm{OPT}} y'_i \cdot p(i) \cdot \Pr[i \notin T]$ satisfies:
$$\sum_{i \in \mathrm{OPT}} y'_i \cdot p(i) \cdot \Pr[i \notin T] = \sum_{i \in \mathrm{OPT}} (1-\alpha') \cdot p(i) = (1-\alpha') \cdot p(\mathrm{OPT}).$$

This implies that the objective value of an optimal solution for Q-LP is at least $(1-\alpha') \cdot p(\mathrm{OPT})$. Hence, $\sum_{i \in \mathrm{OPT}} y^*_i \cdot p(i) \cdot \Pr[i \notin T] \geq (1-\alpha') \cdot p(\mathrm{OPT})$. ◁

▷ Claim 18. The subset Q satisfies the following properties.

1. $\mathbb{E}\left[g_t(Q \setminus T)\right] \leq (1-\alpha')g_t(\mathrm{OPT}) + k \cdot \psi$, for every $t \in \{1, \ldots, k\}$.
2. $\mathbb{E}\left[p(Q \setminus T)\right] \geq (1-\alpha')p(\mathrm{OPT})$.

Proof. The basic optimal solution $\mathbf{y}^*$ has at least $|\mathrm{OPT}|$ tight constraints. Therefore, at least $|\mathrm{OPT}| - k$ constraints of the form $y_i \geq 0$ or $y_i \leq 1$ are tight, i.e., we have at least $|\mathrm{OPT}| - k$ variables y_i with tight constraint. Let $B = \{i \in \mathrm{OPT} \mid 0 < y^*_i < 1\}$ the set of fractional variables, then $|B| \leq k$. For every $t \in \{1, \ldots, k\}$, the following holds:
$$\begin{aligned}
\mathbb{E}\left[g_t(Q \setminus T)\right] &= \sum_{i \in Q} 1 \cdot \Pr[i \notin T] \cdot g_t(i) \\
&= \sum_{i \in B} \Pr[i \notin T] \cdot g_t(i) + \sum_{i \in Q \setminus B} y^*_i \cdot \Pr[i \notin T] \cdot g_t(i) \\
&\leq k \cdot \psi + \sum_{i \in Q \setminus B} y^*_i \cdot \Pr[i \notin T] \cdot g_t(i) \\
&\leq (1-\alpha') \cdot g_t(\mathrm{OPT}) + k \cdot \psi.
\end{aligned}$$

The second equality holds since $y^*_i = 1$, for every $i \in Q \setminus B$. The first inequality holds since $C = \{i\} \in \mathcal{C}(\mathcal{J})$ is a configuration, for every $i \in B$. Therefore, $g_t(C) \leq \psi$, and $|B| \leq k$. The second inequality follows from the constraints of Q-LP. Furthermore,
$$\mathbb{E}\left[p(Q \setminus T)\right] = \sum_{i \in Q} 1 \cdot p(i) \cdot \Pr[i \notin T] \quad \geq \sum_{i \in Q} y^*_i \cdot p(i) \cdot \Pr[i \notin T] \geq (1-\alpha') \cdot p(\mathrm{OPT}).$$

The first inequality holds since $y^*_i \leq 1$, for every $i \in Q$. The second inequality follows from Claim 17. ◁

We now show that the set Q defined in (6) satisfies properties 1. and 2. in the lemma.

▷ Claim 19. For every $t \in \{1, \ldots, k\}$,
$$\Pr\left[g_t\left(Q \setminus T\right) \geq (1-\alpha') \cdot g_t(\mathrm{OPT}) + k \cdot \psi + \varepsilon^{10} \cdot m\right] \leq \exp\left(-\frac{\varepsilon^{21} \cdot m}{\psi^2}\right).$$

Proof. Let $t \in \{1, \ldots, k\}$. Define $q : \mathcal{C}(\mathcal{I}) \to \mathcal{C}(\mathcal{I}_{\mathrm{OPT}})$ by $q(C) = C \cap Q$ for every $C \in \mathcal{C}(\mathcal{I})$. Also, define $f : \mathcal{C}^\ell \to \mathbb{R}$ by

$$f(C_1, \ldots, C_\ell) = \frac{g_t\left(\bigcup_{r \in \{1,\ldots,\ell\}} q(C_r)\right)}{2 \cdot \psi}$$

for every $(C_1, \ldots, C_\ell) \in \mathcal{C}^\ell$, where $\mathcal{C} = \mathcal{C}(\mathcal{I})$. Since $q(C_r) \in \mathcal{C}(\mathcal{I}_{\mathrm{OPT}})$, for every $r \in \{1, \ldots, \ell\}$, by Lemma 4, there exists $C_1, C_2 \in \mathcal{J}$, such that $C_1 \cup C_2 = q(C_r)$. Thus, $2 \cdot \psi \geq g_t(C_1) + g_t(C_2) \geq g_t(q(C))$. By Lemma 12, f is self-bounding function. We note that

$$\mathbb{E}\left[g_t(T \cap Q)\right] \leq \mathbb{E}\left[g_t(\mathrm{OPT})\right] \leq 2 \cdot m \cdot \psi.$$

The first inequality holds since $T \cap Q \subseteq Q \subseteq \mathrm{OPT}$. By Lemma 4, there exist $2 \cdot m$ configurations in $\mathcal{C}(\mathcal{J})$, whose union is OPT, and each configuration $C \in \mathcal{C}(\mathcal{J})$ satisfies $g_t(C) \leq \psi$; thus, the second inequality holds. Hence, we have

$$\begin{aligned}
&\Pr\left[g_t(Q \setminus T) \geq (1 - \alpha') \cdot g_t(\mathrm{OPT}) + k \cdot \psi + \varepsilon^{10} \cdot m\right] \\
&\quad \leq \Pr\left[g_t(Q \setminus T) \geq \mathbb{E}\left[g_t(Q \setminus T)\right] + \varepsilon^{10} \cdot m\right] \\
&\quad = \Pr\left[\frac{g_t(T \cap Q)}{2 \cdot \psi} \leq \mathbb{E}\left[\frac{g_t(T \cap Q)}{2 \cdot \psi}\right] - \frac{\varepsilon^{10} \cdot m}{2 \cdot \psi}\right] \\
&\quad = \Pr\left[f(R_1, \ldots, R_\ell) \leq \mathbb{E}\left[f(R_1, \ldots, R_\ell)\right] - \frac{\varepsilon^{10} \cdot m}{2 \cdot \psi}\right].
\end{aligned}$$

The first inequality holds by Claim 18. The first equality holds since $g(Q \setminus T) = g(Q) - g(Q \cap T)$. Thus,

$$\begin{aligned}
&\Pr\left[g_t(Q \setminus T) \geq (1 - \alpha') \cdot g_t(\mathrm{OPT}) + k \cdot \psi + \varepsilon^{10} \cdot m\right] \\
&\quad \leq \Pr\left[f(R_1, \ldots, R_\ell) \leq \mathbb{E}\left[f(R_1, \ldots, R_\ell)\right] - \frac{\varepsilon^{10} \cdot m}{2 \cdot \psi}\right] \\
&\quad \leq \exp\left(-\frac{\left(\frac{\varepsilon^{10} \cdot m}{2 \cdot \psi}\right)^2}{2 \cdot \mathbb{E}[f(R_1, \ldots, R_\ell)]}\right) \leq \exp\left(-\frac{\varepsilon^{20} \cdot m^2}{2 \cdot 2m \cdot 4 \cdot \psi^2}\right) \leq \exp\left(-\frac{\varepsilon^{21} \cdot m}{\psi^2}\right).
\end{aligned}$$

For the first inequality we used Lemma 11 with $t = \frac{\varepsilon^{10} \cdot m}{2 \cdot \psi}$. The second inequality holds since $\mathbb{E}\left[f(R_1, \ldots, R_\ell)\right] = \frac{\mathbb{E}[g_t(T \cap Q)]}{2 \cdot \psi} \leq 2 \cdot m$. The third inequality holds since $\varepsilon \cdot 16 \leq 1$. ◁

▷ Claim 20. $\Pr\left[p(Q \setminus T) \leq (1 - \alpha' - \varepsilon) \cdot p(\mathrm{OPT})\right] \leq \exp\left(-\varepsilon^{-7}\right)$.

Proof of Claim 20. Let $\tilde{p} : I \to \mathbb{R}_{\geq 0}$ such that $\tilde{p}(i) = p(i)$ for $i \in Q$, and $\tilde{p}(i) = 0$ for $i \notin Q$. Define $f : \mathcal{C}^\ell \to \mathbb{R}$ by

$$f(C_1, \ldots, C_\ell) = \frac{\tilde{p}(\bigcup_{i \in [\ell]} C_i)}{\varepsilon^{10} \cdot p(\mathrm{OPT})}$$

for every $(C_1, \ldots, C_\ell) \in \mathcal{C}^\ell$, where $\mathcal{C} = \mathcal{C}(\mathcal{I})$. Since the instance $\mathcal{I}$ is ε-*nice*,

$$\varepsilon^{10} \cdot p(\mathrm{OPT}) \geq \max_{C \in \mathcal{C}(\mathcal{I})} p(C) \geq \max_{C \in \mathcal{C}(\mathcal{I})} \tilde{p}(C).$$

Therefore, by Lemma 12, f is self bounding function. Now, we can use Lemma 11 and get the following.

$$\begin{aligned}
&\Pr\left[p(Q\setminus T)\leq(1-\alpha')\cdot p(\mathrm{OPT})-\varepsilon\cdot p(\mathrm{OPT})\right]\\
&\quad\leq \Pr\left[p(Q\setminus T)\leq \mathbb{E}\left[p(Q\setminus T)\right]-\varepsilon\cdot p(\mathrm{OPT})\right]\\
&\quad= \Pr\left[\tilde{p}(T)\geq \mathbb{E}\left[\tilde{p}(T)\right]+\varepsilon\cdot p(\mathrm{OPT})\right]\\
&\quad= \Pr\left[\frac{\tilde{p}(T)}{\varepsilon^{10}\cdot p(\mathrm{OPT})}\geq \mathbb{E}\left[\frac{\tilde{p}(T)}{\varepsilon^{10}\cdot p(\mathrm{OPT})}\right]+\varepsilon^{-9}\right]\\
&\quad= \Pr\left[f(R_1,\ldots,R_\ell)\geq \mathbb{E}\left[f(R_1,\ldots,R_\ell)\right]+\varepsilon^{-9}\right].
\end{aligned}$$

The first inequality holds by Claim 18. The first equality follows from subtracting $p(Q)$ for both sides and using $p(Q\setminus T)=p(Q)-p(Q\cap T)$. Thus,

$$\begin{aligned}
&\Pr\left[p(Q\setminus T)\leq(1-\alpha')\cdot p(\mathrm{OPT})-\varepsilon\cdot p(\mathrm{OPT})\right]\\
&\quad\leq \Pr\left[f(R_1,\ldots,R_\ell)\geq \mathbb{E}\left[f(R_1,\ldots,R_\ell)\right]+\varepsilon^{-9}\right]\\
&\quad\leq \exp\left(-\frac{\varepsilon^{-18}}{2\cdot\mathbb{E}[f(R_1,\ldots,R_\ell)]+\frac{\varepsilon^{-9}}{3}}\right)\\
&\quad\leq \exp\left(-\frac{\varepsilon^{-18}}{2\cdot\frac{p(\mathrm{OPT})}{\varepsilon^{10}\cdot p(\mathrm{OPT})}+\frac{\varepsilon^{-9}}{3}}\right)\\
&\quad\leq \exp\left(-\varepsilon^{-7}\right).
\end{aligned}$$

The first inequality follows from using Lemma 11 with $t=\varepsilon^{-9}$. In the second inequality we used the inequality $\mathbb{E}\left[\frac{\tilde{p}(T)}{\varepsilon^{10}\cdot p(\mathrm{OPT})}\right]\leq\frac{p(\mathrm{OPT})}{\varepsilon^{10}\cdot p(\mathrm{OPT})}$. The third inequality holds since $\varepsilon^{-10}\geq\frac{\varepsilon^{-9}}{3}$ and $\varepsilon\cdot 3\leq 1$. ◁

By Claim 19 and Claim 20, we have the statement of the lemma. ◀

Proof of Lemma 8. Let Q be the set defined in Lemma 16, and let F_t be the event "$g_t(Q\setminus T)\leq(1-\alpha')\cdot g_t(S)+k\cdot\psi+\varepsilon^{10}\cdot m$", for every $t\in\{1,\ldots,k\}$. Also, let F_p be the event "$p(Q\setminus T)\geq(1-\alpha'-\varepsilon)\cdot p(\mathrm{OPT})$". By Lemma 16, the following holds:

1. $\Pr[F_t]\geq 1-\exp(-\frac{\varepsilon^{21}\cdot m}{\psi^2})$, for every $t\in\{1,\ldots,k\}$.
2. $\Pr[F_p]\geq 1-\exp(-\varepsilon^{-7})$.

Therefore, by the union bound, we have

$$\begin{aligned}
\Pr\left[F_p\cap\bigcap_{t\in\{0,\ldots,k\}}F_t\right]&\geq 1-\left(\exp(-\varepsilon^{-7})+\sum_{t\in\{0,\ldots,k\}}\exp\left(-\frac{\varepsilon^{21}\cdot m}{\psi^2}\right)\right)\\
&=1-\exp(-\varepsilon^{-7})-k\cdot\exp\left(-\frac{\varepsilon^{21}\cdot m}{\psi^2}\right)\geq\frac{3}{4}.
\end{aligned}$$

The second inequality holds since $k\leq\exp^{[3]}(\varepsilon^{-2})$, $\psi\leq 1$, and $m\geq\exp^{[3]}(\varepsilon^{-30})$.

▷ Claim 21. Assuming that $F_p\cap\bigcap_{t\in\{1,\ldots,k\}}F_t$ occurs,

$$\texttt{BP-OPT}(Q\setminus T,w')\leq\left(2\cdot(1-\alpha')+6\cdot\varepsilon^2\right)\cdot m.$$

Proof. We note that

$$\begin{aligned}
\max_{t\in\{1,\ldots,k\}}g_t(Q\setminus T)&\leq(1-\alpha')\cdot\max_{t\in\{1,\ldots,k\}}\{g_t(\mathrm{OPT})\}+k\cdot\psi+\varepsilon^{10}\cdot m\\
&\leq(1-\alpha')\cdot\texttt{BP-OPT}(\mathrm{OPT},w')+k\cdot\psi+\varepsilon^{10}\cdot m\\
&\leq\left(2\cdot(1-\alpha')+\varepsilon^{10}\right)\cdot m+k\cdot\psi.
\end{aligned}$$

The first inequality holds since $\bigcap_{t\in\{1,\dots,k\}} F_t$ occurs. The second inequality holds since $g_1,\dots,g_k$ are the $(1+\varepsilon^2,\varepsilon^2)$-subset oblivious functions of (OPT,w'). The third inequality holds since there exist m configurations in $\mathcal{C}(\mathcal{I})$ whose union is OPT. Therefore, by Lemma 4, there exist in $\mathcal{C}(\mathcal{J})$ $2\cdot m$ configurations whose union is OPT. Thus, $\texttt{BP-OPT}(\mathrm{OPT},w')\leq 2\cdot m$. We have that

$$\begin{aligned}\texttt{BP-OPT}(Q\setminus T,w') &\leq (1+\varepsilon^2)\cdot \max_{t\in\{1,\dots,k\}}\{g_t(Q\setminus T)\}+\varepsilon^2\cdot\texttt{BP-OPT}(\mathrm{OPT},w')+\delta\\ &\leq (1+\varepsilon^2)\cdot\left(\left(2\cdot(1-\alpha')+\varepsilon^{10}\right)\cdot m+k\cdot\psi\right)+\varepsilon^2\cdot 2\cdot m+\delta\\ &\leq \left(2\cdot(1-\alpha')+5\cdot\varepsilon^2\right)\cdot m+2\cdot k\cdot\psi+\delta\\ &\leq \left(2\cdot(1-\alpha')+5\cdot\varepsilon^2\right)\cdot m+2\cdot\exp^{[3]}(\varepsilon^{-2})+\frac{4}{\varepsilon^8}\\ &\leq \left(2\cdot(1-\alpha')+6\cdot\varepsilon^2\right)\cdot m.\end{aligned}$$

The first inequality follows from Definition 14. The second inequality follows from Lemma 4. The fourth inequality holds since $k\leq\exp^{[3]}(\varepsilon^{-2})$ and $\delta\leq\frac{4}{\varepsilon^8}$. The fifth inequality holds since $\varepsilon^2\cdot m\geq 2\cdot\exp^{[3]}(\varepsilon^{-2})+\frac{4}{\varepsilon^8}$. ◁

By Claim 21, there exist $\left(2\cdot(1-\alpha')+6\cdot\varepsilon^2\right)\cdot m$ configurations in $\mathcal{C}(\mathcal{J})$ whose union is $Q\setminus T$. Among these configurations, we choose the $(1-\alpha')\cdot m$ most profitable. Let R be the items chosen in these configurations. Then,

$$\begin{aligned}p(R) &\geq \frac{(1-\alpha')\cdot m}{(2\cdot(1-\alpha')+6\cdot\varepsilon^2)\cdot m}\cdot p(Q\setminus T)\\ &\geq \frac{(1-\alpha')\cdot m}{(2\cdot(1-\alpha')+6\cdot\varepsilon^2)\cdot m}\cdot(1-\alpha'-\varepsilon)\cdot p(\mathrm{OPT})\\ &= \frac{1-\alpha'}{2\cdot(1-\alpha')+6\cdot\varepsilon^2}\cdot(1-\alpha'-\varepsilon)\cdot p(\mathrm{OPT})\\ &\geq \frac{1-\alpha'-\varepsilon}{2+\varepsilon}\cdot p(\mathrm{OPT})\geq\left(\frac{1-\alpha-2\varepsilon}{2}-\frac{\varepsilon}{4}\right)\cdot p(\mathrm{OPT})\geq\left(\frac{1-\alpha}{2}-2\cdot\varepsilon\right)\cdot p(\mathrm{OPT}).\end{aligned}$$

The third inequality holds since $\frac{6\cdot\varepsilon^2}{1-\alpha'}\leq\varepsilon$. The fourth inequality follows from

$$\frac{1-\alpha'-\varepsilon}{2+\varepsilon}\geq\frac{1-\alpha'-\varepsilon}{2}-\frac{(1-\alpha'-\varepsilon)\cdot\varepsilon}{4}\geq\frac{1-\alpha'-\varepsilon}{2}-\frac{\varepsilon}{4},$$

and by using $\alpha'\leq\alpha+\varepsilon$. Therefore, with probability at least $\frac{3}{4}$, the optimal profit of the MK instance $(Q\setminus T,w',p,(1-\alpha')\cdot m)$ is at least $p(R)$. Since $Q\setminus T\subseteq I\setminus T$, the optimal profit of the MK instance $(I\setminus T,w',p,(1-\alpha')\cdot m)$ is at least $p(R)$. As we obtain a $1-\varepsilon$-approximation for $\mathcal{I}'$ (using, e.g., [10]), the profit of Step 4 in Algorithm 1 is at least

$$(1-\varepsilon)\cdot p(R)\geq\left(\frac{1-\alpha}{2}-2\cdot\varepsilon\right)\cdot(1-\varepsilon)\cdot p(\mathrm{OPT})\quad\geq\left(\frac{1-\alpha}{2}-3\cdot\varepsilon\right)\cdot p(\mathrm{OPT}).$$ ◀

4 APX-hardness

In this section we prove Theorem 1. We use a reduction from 2VBP to show that a PTAS for 2VMK would imply the existence of an APTAS for 2VBP. We rely on the following result of Ray [17], which addresses a flaw in an earlier proof of Woeginger [21].

▶ **Theorem 22** ([17]). *Assuming* P $\neq$ NP, *there is no* APTAS *for* 2VBP.

Proof of Theorem 1. Assume towards a contradiction that there is a PTAS $\{\mathcal{A}_\varepsilon\}$ for 2VMK. That is, for every $\varepsilon > 0$, $\mathcal{A}_\varepsilon$ is a polynomial-time $(1-\varepsilon)$-approximation algorithm for 2VMK.

In Algorithm 2 we use $\{A_\varepsilon\}$ to derive an APTAS for 2VBP. The algorithm calls as a subroutine algorithm First-Fit (FF). The input for FF is an instance $\mathcal{I}$ of 2VMK. The output is a feasible packing of all items in $\mathcal{I}$ in a set of 2-dimensional bins with unit size in each dimension. First-Fit proceeds by considering the items in arbitrary order and assigning the next item in the list to the first bin which can accommodate the item. If no such bin exists, FF opens a new bin and assigns the item to this bin. (For more details on FF and its analysis see, e.g., [19].) Given a 2VBP instance, we use in Algorithm 2 the notion of associated 2VMK instance.

▶ **Definition 23.** *Given a* 2VBP *instance* $\mathcal{I} = (I, w)$ *and* $t \geq 1$*, we define its* t*-associated* 2VMK *instance* $\mathcal{I}' = (I', w', p', m')$ *as follows. The set of items is* $I' = I$*, and the weight function is* $w' = w$*. The profit of each item* $i \in I'$ *is* $p'(i) = w_1(i) + w_2(i)$*, and the number of bins is* $m' = t$*.*

Algorithm 2 Reduction from 2VBP.

configuration : A PTAS $\{\mathcal{A}_\varepsilon\}$ for 2VMK and $\varepsilon > 0$
input : A 2VBP instance $\mathcal{I} = (I, w)$.
output : A solution for $\mathcal{I}$ which uses at most $(1+\varepsilon) \cdot \texttt{BP-OPT}(\mathcal{I}) + 2$ bins.

1 **for** $t = 1$ *to* $|I|$ **do**
2   Let $\mathcal{I}' = (I', w', p', m')$ be the t-associated 2VMK instance of $\mathcal{I}$
3   Use $\mathcal{A}_{\varepsilon'}$ to solve $\mathcal{I}'$, where $\varepsilon' = \frac{\varepsilon}{16}$. Let $C_1, \ldots, C_{m'} \subseteq I$ be the returned solution.
4   Define $S = \bigcup_{j=1}^{m'} C_j$ and pack the residual items $I \setminus S$ using First-Fit.
5   Add $C_1, \ldots, C_m$ along with the packing of S to a list of candidate solutions.
6 **end**
7 Return the candidate solution with the smallest number of bins used.

▷ Claim 24. For any 2VBP instance $\mathcal{I}$ and $\varepsilon > 0$, Algorithm 2 returns a solution which uses at most $(1+\varepsilon) \cdot \texttt{BP-OPT}(\mathcal{I}) + 2$ bins.

Proof. Let $R = \texttt{BP-OPT}(\mathcal{I})$ and let $\mathcal{I}' = (I', w', p', m')$ be the R-associated instance of $\mathcal{I}$. Note that $\text{OPT}(\mathcal{I}') = \sum_{i \in I} p'(i)$, since we can pack all the items I in R bins with the weight function w. Consider the iteration of Step 1 in Algorithm 2 where $t = R$. We show that the number of bins used by the solution is at most $(1+\varepsilon) \cdot \texttt{BP-OPT}(I, w) + 2$. Observe that S, the set in Step 4 of Algorithm 2, satisfies $p'(S) \geq (1 - \frac{\varepsilon}{16}) \cdot \text{OPT}(\mathcal{I}')$. Thus, $p'(I \setminus S) \leq \frac{\varepsilon}{16} \cdot \text{OPT}(\mathcal{I}')$. It follows that

$$\begin{aligned} w_1(I \setminus S) + w_2(I \setminus S) &\leq \frac{\varepsilon}{16} \cdot \text{OPT}(\mathcal{I}') \\ &= \frac{\varepsilon}{16} \cdot (w_1(I) + w_2(I)) \\ &\leq \frac{\varepsilon}{16} \cdot 2 \max\{w_1(I), w_2(I)\} \\ &\leq \frac{\varepsilon \cdot \texttt{BP-OPT}(I, w)}{8}. \end{aligned}$$

The third inequality holds since every bin $b \in \mathcal{C}(\mathcal{I}')$ satisfies $w_1(b), w_2(b) \leq 1$. In the packing of the residual items $I \setminus S$ using First-Fit, every two consecutive bins b_1, b_2 satisfy

$w_1(b_1 \cup b_2) + w_2(b_1 \cup b_2) > 1$. Assume the items in $I \setminus S$ are packed in at least $\varepsilon \cdot \texttt{BP-OPT}(I, w) + 2$ bins in Step 4 of Algorithm 2. Then,

$$w_1(I \setminus S) + w_2(I \setminus S) \geq \left\lfloor \frac{\varepsilon \cdot \texttt{BP-OPT}(I, w) + 2}{2} \right\rfloor > \frac{\varepsilon \cdot \texttt{BP-OPT}(I, w) + 2}{2} - 1 \geq w_1(I \setminus S) + w_2(I \setminus S)$$

The first inequality holds since there are at least $\left\lfloor \frac{\varepsilon \cdot \texttt{BP-OPT}(I,w)+2}{2} \right\rfloor$ disjoint pairs of consecutive bins, and each pair b_1, b_2 satisfies $w_1(b_1 \cup b_2) + w_2(b_1 \cup b_2) \geq 1$. This is a contradiction. Therefore, at most $\varepsilon \cdot \texttt{BP-OPT}(I, w) + 2$ bins are used for the items $I \setminus S$. Hence, Algorithm 2 returns a solution with at most $(1 + \varepsilon) \cdot \texttt{BP-OPT}(\mathcal{I}, w) + 2$ bins. ◁

For every constant $\varepsilon > 0$, it also holds that Algorithm 2 runs in polynomial time. Thus, by Claim 24, it follows that Algorithm 2 is an APTAS for 2VBP. ◀

5 Concluding Remarks

In this paper we present a randomized $\left(1 - \frac{\ln 2}{2} - \varepsilon\right) \approx 0.653$-approximation algorithm for 2VMK, for every fixed $\varepsilon > 0$, thus improving the ratio of $(1 - e^{-1} - \varepsilon) \approx 0.632$, which follows from the results of [8] for the SEPARABLE ASSIGNMENT PROBLEM. To the best of our knowledge, this work is the first direct study of 2VMK in the arena of approximation algorithms.

As an interesting direction for future work, we note that our approach, which combines a technique of [8] with a solution for a residual (1-dimensional) MK instance, does not scale to higher dimensions. Specifically, for dVMK instances where $d \geq 3$, the residual algorithm will obtain marginal profit of $\frac{1}{d \cdot m} \cdot \text{OPT}(\mathcal{I})$ per configuration, which is always lower than the marginal profit obtained by the randomized rounding of [8], due to (4). Hence, for $d \geq 3$, a better approximation ratio is achieved by random sampling of the whole solution as in [8]. We believe that this bottleneck can be resolved by an iterative randomized rounding approach, similar to the approach used in [13] for 2-DIMENSIONAL VECTOR BIN PACKING. This approach can potentially lead also to an improved approximation for 2VMK.

References

1 Max Bannach, Sebastian Berndt, Marten Maack, Matthias Mnich, Alexandra Lassota, Malin Rau, and Malte Skambath. Solving Packing Problems with Few Small Items Using Rainbow Matchings. In *Proc. of MFCS*, pages 11:1–11:14, 2020.

2 Nikhil Bansal, Alberto Caprara, and Maxim. Sviridenko. A new approximation method for set covering problems, with applications to multidimensional bin packing. *SIAM Journal on Computing*, pages 1256–1278, 2010.

3 Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. A sharp concentration inequality with applications. *Random Structures & Algorithms*, 16(3):277–292, 2000.

4 Valentina Cacchiani, Manuel Iori, Alberto Locatelli, and Silvano Martello. Knapsack problems-an overview of recent advances. part ii: Multiple, multidimensional, and quadratic knapsack problems. *Computers & Operations Research*, 2022.

5 Ricardo Stegh Camati, Alcides Calsavara, and Luiz Lima Jr. Solving the virtual machine placement problem as a multiple multidimensional knapsack problem. *ICN 2014*, 264, 2014.

6 Chandra Chekuri and Sanjeev Khanna. A polynomial time approximation scheme for the multiple knapsack problem. *SIAM Journal on Computing*, 35(3):713–728, 2005.

7 Tome Cohen, Ariel Kulik, and Hadas Shachnai. Improved approximation for two-dimensional vector multiple knapsack. *arXiv preprint*, 2023. `arXiv:2307.02137`.

8 Lisa Fleischer, Michel X Goemans, Vahab S Mirrokni, and Maxim Sviridenko. Tight approximation algorithms for maximum separable assignment problems. *Mathematics of Operations Research*, 36(3):416–431, 2011.

9 Alan M Frieze, Michael RB Clarke, et al. Approximation algorithms for the m-dimensional 0-1 knapsack problem: worst-case and probabilistic analyses. *European Journal of Operational Research*, 15(1):100–109, 1984.

10 Klaus Jansen. Parameterized approximation scheme for the multiple knapsack problem. *SIAM Journal on Computing*, 39(4):1392–1412, 2010.

11 Klaus Jansen. A fast approximation scheme for the multiple knapsack problem. In *International Conference on Current Trends in Theory and Practice of Computer Science*, pages 313–324, 2012.

12 Hans Kellerer, Ulrich Pferschy, and David Pisinger. *Knapsack problems*. Springer, 2004.

13 Ariel Kulik, Matthias Mnich, and Hadas Shachnai. Improved approximations for vector bin packing via iterative randomized rounding. In *Proc. of FOCS (to appear)*, 2023.

14 Ariel Kulik and Hadas Shachnai. There is no EPTAS for two-dimensional knapsack. *Information Processing Letters*, 110(16):707–710, 2010.

15 Alexandra Lassota, Aleksander Łukasiewicz, and Adam Polak. Tight Vector Bin Packing with Few Small Items via Fast Exact Matching in Multigraphs. In *Proc. of ICALP*, pages 87:1–87:15, 2022.

16 Colin McDiarmid et al. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

17 Arka Ray. There is no APTAS for 2-dimensional vector bin packing: Revisited. *arXiv preprint*, 2021. `arXiv:2104.13362`.

18 Yang Song, Chi Zhang, and Yuguang Fang. Multiple multidimensional knapsack problem and its applications in cognitive radio networks. In *MILCOM 2008-2008 IEEE Military Communications Conference*, pages 1–7. IEEE, 2008.

19 Vijay V Vazirani. *Approximation algorithms*, volume 1. Springer, 2001.

20 Jan Vondrák. A note on concentration of submodular functions. *arXiv preprint*, 2010. `arXiv:1005.2791`.

21 Gerhard J Woeginger. There is no asymptotic PTAS for two-dimensional vector packing. *Information Processing Letters*, 64(6):293–297, 1997.