Approximating robust bin-packing with budgeted uncertainty

Aniket Basu Roy, Marin Bougeret, Noam Goldberg, Michael Poss

Approximating robust bin-packing with budgeted uncertainty

Aniket Basu Roy[1], Marin Bougeret[1], Noam Goldberg[2], and Michael Poss[1]

[1] LIRMM, University of Montpellier, CNRS, Montpellier, France
{aniket.basu-roy,marin.bougeret,michael.poss}@lirmm.fr
[2] Department of Management, Bar-Ilan University, Ramat Gan 5290002, Israel
noam.goldberg@biu.ac.il

Abstract. We consider robust variants of the bin-packing problem where the sizes of the items can take any value in a given uncertainty set $U \subseteq \times_{i=1}^{n}[\overline{a}_i, \overline{a}_i + \hat{a}_i]$, where $\overline{a} \in [0,1]^n$ represents the nominal sizes of the items and $\hat{a} \in [0,1]^n$ their possible deviations. We consider more specifically two uncertainty sets previously studied in the literature. The first set, denoted U^{Γ}, contains scenarios in which at most $\Gamma \in \mathbb{N}$ items deviate, each of them reaching its peak value $\overline{a}_i + \hat{a}_i$, while each other item has its nominal value $\overline{a}_i$. The second set, denoted U^{Ω}, bounds by $\Omega \in [0,1]$ the total amount of deviation in each scenario. We show that a variant of the next-fit algorithm provides a 2-approximation for model U^{Ω}, and a $2(\Gamma+1)$ approximation for model U^{Γ} (which can be improved to 2 approximation for $\Gamma = 1$). This motivates the question of the existence of a constant ratio approximation algorithm for the U^{Γ} model. Our main result is to answer positively to this question by providing a 4.5 approximation for U^{Γ} model based on dynamic programming.

Keywords: Bin-packing · Robust optimization · Approximation Algorithm · Next-fit · Dynamic programming

1 Introduction

Bin packing is the problem of assigning a given set of n items, each item of a specified size, to the smallest number of unit capacity bins. The problem has been the subject of study in an extensive body of research initiated by several publications in the 1970s including the work of Johnson et al. [10]. The problem is $\mathcal{NP}$-hard and in fact a straightforward reduction from the partition decision problem implies that it is $\mathcal{NP}$-hard to determine whether a bin-packing instance has a solution using only two bins. This also shows that the problem cannot be approximated within a factor less than 3/2. An approximation factor guarantee of 3/2 has been proven for the first-fit decreasing algorithm by Simchi-Levi [15]. Much of the research has concentrated on the asymptotic setting where n tends

This research has benefited from the support of the ANR project ROBUST [ANR-16-CE40-0018].

to infinity, and in the online setting where the instance is not given in advance but each item is revealed and packed one at a time. A fully polynomial-time approximation scheme for the offline asymptotic problem is due to Karmarkar and Karp [11]. The best asymptotic and absolute online competitive ratios of 1.578 and 5/3, respectively, are due to Balogh et al.in [4] and [3], respectively.

In many applications, the sizes of the items to be packed are not fully known at the time that the packing is carried out. In cargo shipping, for example, the actual weight of a container may deviate from its declared weight or its measurements may be inaccurate. Bin packing has been used to model the assignment of elective surgeries to operating room in hospitals [7]. Here a bin is a shift of a properly equipped and staffed operating room for performing a certain type of elective surgeries. The room scheduler has to fit in the bins as many cases (patients) as possible. In this setting clearly the length of time of performing each surgery is subject to uncertainty for example in the event of complications. One way to model the uncertainty that falls into the framework of robust optimization is to assume that the sizes are uncertain parameters taking any value in a given set $U \subset \mathbb{R}^n$, where each $a \in U$ represents a possible scenario. This leads to the following problem (where the description of U is sometimes not explicit to avoid exponential length in n)

RBP (Robust bin-packing)
Input: $U \subset \mathbb{R}^n$
Output: A solution is a partition of $[n]$ into k bins $b_1, \ldots, b_k$ such that $\max_{a \in U} \sum_{i \in b_j} a_i \leq 1$ for each $j \in [k]$
Minimize: k

Classically, robust combinatorial optimization has dealt with uncertain objective, meaning that the cost vector c can take any value in set U, unlike RBP where the uncertainty affects the feasibility of the solutions. In that context, it is well-known that arbitrary uncertainty sets U lead to robust counterparts that are hardly approximable. For instance, the robust knapsack is not approximable at all [1], while the shortest path, the spanning tree, the minimum cut, and the assignment problem do not admit constant-ratios approximation algorithms, e.g. [12, 13]. Furthermore, describing U by an explicit list of scenarios runs the risk of over-fitting so the optimal solutions may become infeasible for small variations outside U. These two drawbacks are usually tackled by using more specific uncertainty sets, defined by simple budget constraints. One of these widely used uncertainty sets, U^Γ, supposes that the size of each item is either its given nominal size $\bar{a}_i$, or its peak value $\bar{a}_i + \hat{a}_i$. Furthermore, in any scenario, at most $\Gamma \in \mathbb{N}$ of the items may assume their peak value simultaneously. Formally, U^Γ can be defined as $U^\Gamma = \{a | \forall i \in [n], a_i \in \{\bar{a}_i, \bar{a}_i + \hat{a}_i\}$ and $\sum_{i \in [n]} (a_i - \bar{a}_i)/\hat{a}_i \leq \Gamma\}$.[3] Set U^Γ has had a large success in robust combinatorial optimization with a constant number of constraints because the set essentially preserves the complexity and

[3] U^Γ is often defined alternatively in the literature, as the polytope $\{a \in \times_{i \in [n]} [\bar{a}_i, \bar{a}_i + \hat{a}_i] \mid \sum_{i \in [n]} (a_i - \bar{a}_i)/\hat{a}_i \leq \Gamma\}$. For the bin-packing problem, one readily verifies using classical arguments that the two definitions lead to the same optimization problem.

approximability properties of the nominal problem. The result was initially proposed for min-max problems in [5], and was independently extended to uncertain constraints in [2, 8], contrasting with the aforementionned uncertain objective. We also consider a second uncertainty set (used in [9, 17], among others), characterized again by $\bar{a}$ and $\hat{a}$, as well as the number $\Omega \in [0,1]$ stating how much deviation can be spread among all sizes, formally $U^{\Omega} = \{a \in \times_{i\in[n]}[\bar{a}_i, \bar{a}_i + \hat{a}_i] \mid \sum_{i\in[n]}(a_i - \bar{a}_i) \leq \Omega\}$. From the approximability viewpoint, set U^{Ω} benefits from similar positive results as U^{Γ}, see [14].

The above positive complexity results (e.g. [8, 14]) imply, for instance, that there exists a fully-polynomial time approximation scheme (FPTAS) for the robust knapsack problem with uncertain profits and uncertain weights belonging to U^{Ω} and/or U^{Γ}. Interestingly, these positive results do not extend to most scheduling problems (because they involve non-linearities) and to the bin-packing problem (because it involves a non-constant numbers of robust constraints). While in a previous paper [6] (with authors in common) we provided approximability results on robust scheduling, no such results have yet been proposed for the bin-packing problem, the only previous work focusing on numerical algorithms [16]. The purpose of this paper is to fill these gaps, as we present constant-ratio approximation algorithms the bin-packing problem, both for U^{Ω} and U^{Γ}.

Notations, problems definitions, and next-fit algorithm In this paper we consider two special cases of RBP. In the first one, ΓRBP, the input is $\mathcal{I} = (n, \overline{a}, \hat{a}, \Gamma)$ where $n \in \mathbb{N}$, and we assume that $U = U^{\Gamma}$. In the second one, ΩRBP, the input is $\mathcal{I} = (n, \overline{a}, \hat{a}, \Omega)$ where $n \in \mathbb{N}$, $\overline{a} \in [0,1]^n$, $\hat{a} \in [0,1]^n$, and $\Omega \in [0,1]$, and we assume that $U = U^{\Omega}$.

Let us now provide some important notations that will allow us to restate ΓRBP and ΩRBP in a more convenient way. Given $n \in \mathbb{N}$, sets $\{0, 1, \ldots, n\}$ and $\{1, \ldots, n\}$ are respectively denoted $[n]_0$ and $[n]$. Set $\{i, \ldots, j\}$ is denoted by $[\![i, j]\!]$. Given a vector $v \in [0,1]^n$ and a subset $X \subseteq [n]$, we define $v(X) = \sum_{i\in X} v_i$. Given two vectors $\overline{a} \in [0,1]^n$, $\hat{a} \in [0,1]^n$ and a subset of items $X \subseteq [n]$, we define $\hat{a}_{\Omega}(X) = \min\{\hat{a}(X), \Omega\}$, $\Gamma(X)$ as the set of Γ items in X with largest $\hat{a}$ values (ties broken by taking smallest indices), or $\Gamma(X) = X$ if $|X| < \Gamma$, and $\hat{a}_{\Gamma}(X) = \hat{a}(\Gamma(X))$. Accordingly, we define the fill of a bin $b \subseteq [n]$ as $f_{\Gamma}(b) = \overline{a}(b) + \hat{a}_{\Gamma}(b)$ for set U^{Γ}, and $f_{\Omega}(b) = \bar{a}(b) + \hat{a}_{\Omega}(b)$ for set U^{Ω}. The fill of a bin for a general uncertainty set U is denoted as $f_U(b) = \max_{a\in U} a(b)$.

Consider the following example. We are given an ordered set of pairs $(\bar{a}_i, \hat{a}_i)$, $X = \{(0.3, 0.2), (0.4, 0.2), (0.3, 0.1), (0.2, 0.5)\}$ with $\Gamma = 2$ and $\Omega = 0.3$. Thus, $\Gamma(X) = \{(0.3, 0.2), (0.2, 0.5)\}$, $\bar{a}(X) = 1.2$, $\hat{a}_{\Gamma}(X) = 0.7$, and $f_{\Gamma}(X) = 1.9$. Similarly, $\hat{a}_{\Omega}(X) = 0.3$ and $f_{\Omega}(X) = 1.0$.

Now, observe that $\max_{a\in U} \sum_{i\in b_j} a_i \leq 1$ (the constraint required in RBP) is equivalent to $f_U(b) \leq 1$, and thus to $f_{\Gamma}(b_j) \leq 1$ for ΓRBP and $f_{\Omega}(b_j) \leq 1$ for ΩRBP. For example in ΓRBP, $f_{\Gamma}(b_j) \leq 1$ simply means that the total nominal ($\overline{a}$) size of the items plus the deviating size ($\hat{a}$) of the Γ largest (in $\hat{a}$ values)

items must not exceed one. Thus, the two optimization problems studied in this paper can be equivalently formulated in the following way.

ΓRBP (Γ-robust bin-packing)
Input: $\mathcal{I} = (n, \overline{a}, \hat{a}, \Gamma)$ where $n \in \mathbb{N}$, $\overline{a} \in [0,1]^n$, $\hat{a} \in [0,1]^n$, and $\Gamma \in \mathbb{N}$.
Output: A solution is a partition of $[n]$ into k bins $b_1, \ldots, b_k$ such that $f_\Gamma(b_j) \leq 1$ for each $j \in [k]$
Minimize: k

ΩRBP (Ω-robust bin-packing)
Input: $\mathcal{I} = (n, \overline{a}, \hat{a}, \Omega)$ where $n \in \mathbb{N}$, $\overline{a} \in [0,1]^n$, $\hat{a} \in [0,1]^n$, and $\Omega \in [0,1]$.
Output: A solution is a partition of $[n]$ into k bins $b_1, \ldots, b_k$ such that $f_\Omega(b_j) \leq 1$ for each $j \in [k]$
Minimize: k

The optimal solution value or cost of either problem is denoted by $\mathsf{OPT}(\mathcal{I}) = k^*$ ($\mathcal{I}$ may be omitted when the instance is clear from the context) and a corresponding optimal solution is denoted by $s^* = \{b_1^*, b_2^*, \ldots, b_{k^*}^*\}$. We introduce in Algorithm 1 a variant of the standard next fit algorithm.

initialization: $j = 1$
1 Pack items (with smaller index first) in b_j until $f_U(b_j) > 1$ or $n \in b_j$. If $n \notin b_j$ then $j \leftarrow j + 1$ and repeat Step 1. Otherwise, $k' \leftarrow j$ proceed to Step 2.
2 Pack the last item of each bin in a new bin: for any j, let $i = \max(b_j)$, $b_j^1 = b_j \setminus \{i\}$, and $b_j^2 = \{i\}$
return : $\bigcup_{j=1}^{k'} \{b_j^1, b_j^2\}$

Algorithm 1: NEXT-FIT($\mathcal{I}$)

Structure of the paper In Sections 2 and 3, we analyze the ratio provided by NEXT-FIT for ΩRBP and ΓRBP, respectively. For ΩRBP, using ordering (1) (non increasing ordering on $\frac{\hat{a}_i}{\overline{a}_i}$) the ratio is equal to 2. For ΓRBP, using ordering (2) (non increasing ordering on $\hat{a}_i$), the ratio is bounded by $2(\Gamma + 1)$ (and can be improved to 2 for $\Gamma = 1$). As Theorem 4 shows that neither ordering (1) or (2) leads to a constant ratio using NEXT-FIT, this raises the question the existence of a constant approximation for ΓRBP. In Section 4 we first review some basic ideas and explain why they are not sufficient. Then, we introduce the key elements necessary to develop our dynamic programming algorithm (DP) in Section 5. The latter gives a ratio of 4.5 for ΓRBP and any $\Gamma \in \mathbb{N}$, which is our main result.

2 Next-fit for ΩRBP

Unlike the classical bin-packing problem, executing NEXT-FIT on arbitrarily ordered items can lead to arbitrarily bad solutions. For example, given ϵ with

$0 < \epsilon \leq \frac{1}{2n}$, consider an instance with $\Omega = 1 - \epsilon$, and items $((2\epsilon, 0), (0, 1 - \epsilon), \ldots, (2\epsilon, 0), (0, 1 - \epsilon))$, where item $i \in [n]$ is denoted by the pair $(\bar{a}_i, \hat{a}_i)$. Using this ordering, Next-Fit will create $n/2$ bins b_j with $f_\Omega(b_j) > 1$ for any $j \in [n]$ (which will be turned into n bins $\{b_j^1, b_j^2\}$), whereas the optimal solution uses 2 bins. This example also illustrates that, unlike in the standard bin-packing, the total size argument no longer apply to the robust counterpart as having $f_\Omega(b_j) > 1$ for any $j \in [n]$ does not imply a large (depending on n) lower bound on the optimal.

Next, we consider an ordering of the items such that

$$\hat{a}_1/\bar{a}_1 \geq \cdots \geq \hat{a}_n/\bar{a}_n. \tag{1}$$

Lemma 1. *Suppose that the items are ordered according to* (1)*. Then* $k' \leq k^*$.

Proof. Consider an optimal solution $b_1^*, \ldots, b_{k^*}^*$ and the subset of optimal bins given by $G^* = \{j \in [k^*] \mid \hat{a}(b_j^*) > \Omega\}$. Let

$$A = \sum_{i \in [n]} (\bar{a}_i + \hat{a}_i) = \sum_{j=1}^{k'} (\bar{a}(b_j) + \hat{a}(b_j)) = \sum_{j=1}^{k^*} \left(\bar{a}(b_j^*) + \hat{a}(b_j^*)\right).$$

Let G denote the first $|G^*|$ bins opened in Step 1 of Next-Fit. If $k' \in G$ then clearly $k' \leq k^*$. Otherwise, it can be observed that for each $l \in G$, $\bar{a}(b_l) > 1 - \Omega$ (as $\bar{a}(b_\ell) + \hat{a}_\Omega(b_\ell) > 1$ and $\hat{a}_\Omega(b_\ell) \leq \Omega$) and $1 - \Omega \geq \max_{j \in G^*} \bar{a}(b_j^*)$ (as $f_\Omega(b_j) \leq 1$). Thus, $\sum_{j \in G} \bar{a}(b_j) > \sum_{j \in G^*} \bar{a}(b_j^*)$ and so by the assumed ordering (1) of the items, following a standard knapsack argument, $\sum_{j \in G} \hat{a}(b_j) > \sum_{j \in G^*} \hat{a}(b_j^*)$. Letting $\bar{G} = [k'] \setminus G$ and $\bar{G}^* = [k^*] \setminus G^*$, it follows that

$$\sum_{j \in \bar{G}} (\bar{a}(b_j) + \hat{a}(b_j)) = A - \sum_{j \in G} (\bar{a}(b_j) + \hat{a}(b_j)) \leq$$
$$A - \sum_{j \in G^*} \left(a(b_j^*) + \hat{a}(b_j^*)\right) = \sum_{j \in \bar{G}^*} \left(\bar{a}(b_j^*) + \hat{a}(b_j^*)\right)$$

(equality may hold throughout if $G^* = \emptyset$). Further, for each $j \in \bar{G} \setminus \{k'\}$, $\bar{a}(b_j) + \hat{a}(b_j) \geq f(b_j) > 1$ and for each $j \in \bar{G}^*$, $\bar{a}(b_j^*) + \hat{a}(b_j^*) \leq 1$. Therefore, $|\bar{G}| \leq \left\lceil \sum_{j \in \bar{G}} (\bar{a}(b_j) + \hat{a}(b_j)) \right\rceil \leq \left\lceil \sum_{j \in \bar{G}^*} \left(\bar{a}(b_j^*) + \hat{a}(b_j^*)\right) \right\rceil \leq |\bar{G}^*|$ and $k' \leq k^*$ as claimed. □

The lemma combined with Step 2 of Next-Fit immediately imply the following theorem.

Theorem 1. *If the items are ordered according to* (1) *then* Next-Fit *is a 2-approximation algorithm for* ΩRBP.

3 Next-fit for ΓRBP

From now on, we focus on problem ΓRBP. Remark first that using an arbitrary ordering leads to arbitrarily bad solutions, considering $\Gamma = 1$ and the same

items $((2\epsilon, 0), (0, 1-\epsilon), \ldots, (2\epsilon, 0), (0, 1-\epsilon))$ as in the previous section. Thus, we consider here an ordering of the items such that

$$\hat{a}_1 \geq \cdots \geq \hat{a}_n. \tag{2}$$

The main result of this Section is the following.

Theorem 2. *If the items are ordered according to* (2) *then* Next-Fit *is a* $2(\Gamma+1)$*-approximation algorithm for* ΓRBP*.*

Proof. Recall that k' is the number of bins used in Step 1 and let $s' = (b_1, \ldots, b_{k'})$ be the bins output at the end of Step 1. Recalling that $i^*_\ell = \max(\Gamma(b^*_\ell))$, we denote the index of the bin that contains i^*_ℓ at the end of Step 1 of Next-Fit by $d(\ell)$. We also define $d(0) = 0$ and $d(k^*+1) = +\infty$. In addition, we define $i_j = \max(\Gamma(b_j))$ for each $j \in [k']$.

Using the same notations as in Section 4, given a solution $s = \{b_j, j \in [k]\}$, we define $P(s) = \{\Gamma(b_j), j \in [k]\}$ as the profile of s and $\tilde{P}_j(s) = \{i \mid i \in \Gamma(b_\ell)$ for some $\ell \in [j]\}$ as all deviating items in the first j bins, and $D(s) = \tilde{P}_k(s)$ the set of all deviating items of s.

The purpose of this section is to prove the following counterpart of Lemma 5 when Γ is larger than 1, that result implying immediately the approximation of $2(\Gamma+1)$.

Lemma 2. *Suppose that the items are ordered according to* (2)*. Then* $k' \leq (\Gamma+1)k^*$*.*

Let $\mathcal{M}_{\ell(\ell+1)} = \{b_j, d(\ell) < j < d(\ell+1)\}$ for each $\ell \in [k^*]_0$, so $s' = (\mathcal{M}_{01}, b_{d(1)}, \mathcal{M}_{12}, \ldots, b_{d(k^*)}, \mathcal{M}_{k^*(k^*+1)})$. We introduce the set of all bins but bins $b_{d(\ell)}$ as $\mathcal{M} = \mathcal{M}_{01} \cup \mathcal{M}_{12} \cup \cdots \cup \mathcal{M}_{k(k+1)}$. Let $\mathcal{M}' \subseteq \mathcal{M}$ be the set of bins that contain only items that do not deviate in s^*, that is, $b \cap D(s^*) = \emptyset$ for each $b \in \mathcal{M}'$. We further define $\mathcal{M}'' = \mathcal{M} \setminus \mathcal{M}'$. In what follows, we bound $|\mathcal{M}'|$ and $|\mathcal{M}''|$ by linear functions of k^*.

Lemma 3. $|\mathcal{M}''| \leq (\Gamma-1)k^*$

Proof. As each bin $b \in \mathcal{M}''$ contains at least one item from $D(s^*) \setminus \bigcup_{j\in[k^*]} i^*_j$, and as $|D(s^*) \setminus \bigcup_{j\in[k^*]} i^*_j| \leq (\Gamma-1)k^*$, the result is immediate. □

The case of $\mathcal{M}'$ is a little more involved.

Lemma 4. $|\mathcal{M}'| \leq k^*$

Proof. Let us define $\mathcal{M}'_{\ell(\ell+1)} = \mathcal{M}_{\ell(\ell+1)} \cap \mathcal{M}'$ for each $\ell \in [k^*]$ and let m_ℓ be the number of bins in $\mathcal{M}'_{\ell(\ell+1)}$ for $\ell \in [k^*]$. We do not define $\mathcal{M}'_{01}$ because for each $b \in \mathcal{M}_{01}$, b contains items with larger $\hat{a}$ than $a_{i^*_1}$, so all items of b deviate in s^* and $\mathcal{M}'_{01}$ is empty. We discuss next what are the possible packing of the items of $\mathcal{M}'$ into the bins $b^*_1, \ldots, b^*_{k^*}$ of s^*. Each bin $b \in \mathcal{M}'_{\ell(\ell+1)}$ must be entirely packed into bins $\{b^*_j, j \in [\ell]\}$ because otherwise the elements of b would deviate in b^*_j, which would contradict the definition of $\mathcal{M}'$. Hence, a necessary condition

for that packing to be feasible is that the size in $\overline{a}$ of the bins in $\bigcup_{\ell\in[j]} \mathcal{M}'_{\ell(\ell+1)}$ does not exceed the available space in bins $\{b^*_j, j \in [\ell]\}$:

$$\sum_{\ell\in[j]} \sum_{b\in\mathcal{M}'_{\ell(\ell+1)}} \overline{a}(b) \leq \sum_{\ell\in[j]} (1 - \hat{a}_\Gamma(b^*_\ell)). \tag{3}$$

Let us define $\bar{b}_\ell = 1 - \Gamma\hat{a}_{i^*_\ell}$ for each $\ell \in [k^*]$. By the ordering of the optimal bins,

$$\bar{b}_1 \leq \cdots \leq \bar{b}_{k^*}. \tag{4}$$

Since all deviating items (in the solution computed by Next-Fit) of a bin $b \in \mathcal{M}'_{\ell(\ell+1)}$ have $\hat{a}$ not greater than $\hat{a}_{i^*_\ell}$, we have for each $\ell \in [k^*]$

$$\overline{a}(b) \geq \bar{b}_\ell, \quad \forall b \in \mathcal{M}'_{\ell(\ell+1)}. \tag{5}$$

Similarly, all deviating items of b^*_ℓ have $\hat{a}$ not smaller than $\hat{a}_{i^*_\ell}$ so for each $\ell \in [k^*]$

$$1 - \hat{a}_\Gamma(b^*_\ell) \leq \bar{b}_\ell. \tag{6}$$

In view of (5) and (6), a necessary condition for (3) to hold is the following set of inequalities

$$\sum_{\ell\in[j]} m_\ell \bar{b}_\ell \leq \sum_{\ell\in[j]} \bar{b}_\ell, \quad \forall j \in [k^*]. \tag{7}$$

The rest of the proof is devoted to proving the following algebraic result. Given a set of non-negative integers $m_1, \ldots, m_{k^*}$ and a set of non-negative reals $\bar{b}_1, \ldots \bar{b}_{k^*}$ that satisfy (4) and (7), then it also holds that

$$\sum_{\ell\in[j]} m_\ell \leq j, \quad \forall j \in [k^*]. \tag{8}$$

We see immediately that the k^*-inequality of (8) leads to the statement of the lemma. We prove below that (8) holds by induction on j. The first inequality of (8) is identical to the first inequality of (7), proving the case $j = 1$. Next, consider

$$\sum_{\ell\in[k]} m_\ell \leq k. \tag{9}$$

assuming that (8) holds for $j \in [k-1]$. If $m_k \leq 1$, then $\sum_{\ell\in[k]} m_\ell \leq \sum_{\ell\in[k-1]} m_\ell + 1 \leq k - 1 + 1 = k$, where the second inequality follows from the induction. Otherwise, we define $\mu = m_k - 1$ and from $b_{k-1} \leq b_k$ and (7), we obtain

$$\sum_{\ell\in[k-2]} m_\ell b_\ell + (m_{k-1} + \mu) b_{k-1} + b_k \leq \sum_{\ell\in[k]} m_\ell b_\ell \leq \sum_{\ell\in[k]} b_\ell$$

Subtracting b_k from the above inequalities, we obtain

$$\sum_{\ell\in[k-2]} m_\ell b_\ell + (m_{k-1} + \mu) b_{k-1} \leq \sum_{\ell\in[k-1]} b_\ell. \tag{10}$$

Defining $m'_\ell = m_\ell$ for $\ell \in [k-2]$ and $m'_{k-1} = m_{k-1} + \mu$, (10) becomes $\sum_{\ell\in[k-1]} m'_\ell b_\ell + m'_{k-1} b_{k-1} \leq \sum_{\ell\in[k-1]} b_\ell$, so that m' satisfies inequality (7) for $j = k-1$. As $m'_\ell = m_\ell$ for $\ell \in [k-2]$, m' also satisfies inequality (7) for $j \in [k-2]$, so by induction, we have

$$\sum_{\ell\in[k-1]} m'_\ell \leq k-1. \tag{11}$$

What is more, by definition of m' it holds that $\sum_{\ell\in[k-1]} m'_\ell = \sum_{\ell\in[k]} m_k - 1$. Plugging the previous inequality into the left-hand side of (11) yields (9), proving the result. □

Therefore, lemma 3 and 4 implies that $|\mathcal{M}| \leq \Gamma k^*$. On adding the bins $\{b_{d(l)}, \ell \in [k^*]\}$, the number of bins at the end of step 1 of the Next-Fit algorithm becomes $(\Gamma+1)k^*$, proving thus Lemma 2. Finally, after step 2 of the algorithm the total number of bins can be at most $2(\Gamma+1)k^*$, concluding the proof of Theorem 2. □

As in the previous section, we obtain the following theorem.

Theorem 3. *If the items are ordered according to* (2) *and* $\Gamma = 1$ *then* Next-Fit *is a 2-approximation algorithm for* ΓRBP.

Proof. The deviating item of bin j in a fixed optimal solution s^* and in the solution of Next-Fit are denoted by $\{i^*_j\} = \Gamma(b^*_j)$ and $\{i_j\} = \Gamma(b_j)$, respectively. We order the bins of s^* such that $i^*_j \geq i^*_{j+1}$.Notice that by definition of Next-Fit and ordering (2) we also have $i_j \geq i_{j+1}$.

Lemma 5. *Suppose that the items are ordered according to* (2) *and that* $\Gamma = 1$. *Then* $k' \leq k^*$.

Proof. Suppose by contradiction that $k' > k^*$. Let $b_1, \ldots, b_{k'}$ be the bins opened at Step 1 of Next-Fit and notice that $f_\Gamma(b_j) = \overline{a}(b_j) + \hat{a}_\Gamma(b_j) > 1$ for each $j \in [k'-1]$, while $\overline{a}(b^*_j) + \hat{a}_\Gamma(b^*_j) \leq 1$ for each $j \in [k^*]$. We prove next by induction on $\ell \in [k^*]$ that

$$\sum_{j=1}^{\ell} \overline{a}(b_j) > \sum_{j=1}^{\ell} \overline{a}(b^*_j). \tag{12}$$

For $\ell = 1$, we have $i_1 = i^*_1 = 1$ and (12) follows immediately from $\hat{a}_\Gamma(b_1) = \hat{a}_\Gamma(b^*_1)$. Suppose now that induction hypothesis is true for $\ell - 1$. By definition of i^*_ℓ and i_ℓ, we know that $[i^*_\ell - 1] \subseteq \bigcup_{j=1}^{\ell-1} b^*_j$ and $[i_\ell - 1] = \bigcup_{j=1}^{\ell-1} b_j$. Using induction hypothesis, we get that $i_\ell \geq i^*_\ell$, and accordingly $\hat{a}_\Gamma(b_\ell) \leq \hat{a}_\Gamma(b^*_\ell)$. As $l \leq k^* < l$, we have $f_\Gamma(b_l) > 1$, leading to $\overline{a}(b_\ell) > \overline{a}(b^*_\ell)$.

Thus, for $l = k^*$ we get $\sum_{j=1}^{k^*} \overline{a}(b_j) > \sum_{j=1}^{k^*} \overline{a}(b^*_j) = \sum_{i\in[n]} \overline{a}_i$, which is impossible. □

□

To complete the analysis, we establish the following lower bound on the ratio of Next-Fit.

Theorem 4. *If the items are ordered according to* (2) *or* (1)*, then the approximation ratio of* Next-Fit *for* ΓRBP *is at least* $\frac{2\Gamma}{3}$.

Proof. Let us define an instance where the ordering (2) can lead to Step 1 of Next-Fit using $k' = \Gamma$ bins while $\mathsf{OPT} = 3$. Every row of the $\Gamma \times \Gamma$ matrix below corresponds to the set of items in a bin (after the Step 1) of Next-Fit algorithm

$$\begin{array}{llll} (\epsilon, 1/\Gamma - \delta_1) & (0, 1/\Gamma - \delta_1) & \dots & (0, 1/\Gamma - \delta_1) \\ \vdots & \vdots & \ddots & \vdots \\ (\epsilon, 1/\Gamma - \delta_\Gamma) & (0, 1/\Gamma - \delta_\Gamma) & \dots & (0, 1/\Gamma - \delta_\Gamma) \end{array} \tag{13}$$

where $\epsilon \leq 1/\Gamma$ and $\delta_1 \leq \dots \leq \delta_\Gamma < \epsilon/\Gamma$. On the one hand, $\epsilon + \Gamma \cdot (1/\Gamma - \delta_l) > 1$ for each $l \in [\Gamma]$, so step 1 of Next-Fit outputs Γ bins. On the other hand, an optimal solution can pack all the items above except the ones in the first column into a single bin because $\Gamma \cdot 1/\Gamma - \delta_1 \leq 1$. Further, the total weight of the first $\Gamma/2$ items of the first column sums up to $\Gamma/2 \cdot (1/\Gamma + \epsilon) - \sum_{l=1}^{\Gamma/2} \delta_l \leq 1 - \sum_{l=1}^{\Gamma/2} \delta_l \leq 1$, and similarly for the last $\Gamma/2$ items, so an optimal solution may pack the first column using two bins. Finally, instance (13) shows that Next-Fit produces a solution $2\Gamma/3$ times worse than the optimal one.

This instance can be adapted to establish a lower bound for the approximation ratio ofNext-Fit when items are ordered according to (1). We consider an example that yields a lower bound on the approximation ratio of Next-Fit in solving ΓRBP when the items are ordered according to (1). For some $c \geq \Gamma^2$, $\epsilon' = \frac{2}{\Gamma(1+c)}$, $\epsilon = \frac{1}{\Gamma(\Gamma^2+\Gamma-1)}$ consider the instance given by the following $\Gamma \times \Gamma$ matrix (similar to the instance in Section 3) with entries corresponding to items $(\bar{a}_i, \hat{a}_i)$ for $i \in [n] = [\Gamma^2]$:

$$\begin{array}{llll} (\epsilon', c\epsilon') & (\epsilon, c\epsilon) & \dots & (\epsilon, c\epsilon) \\ \vdots & \vdots & \ddots & \vdots \\ (\epsilon', c\epsilon') & (\epsilon, c\epsilon) & \dots & (\epsilon, c\epsilon) \end{array}$$

It can be verified that Next-Fit opens a bin for each row, since $(1+c)\epsilon' + (\Gamma - 1)(1+c)\epsilon > 1$. The optimal solution opens 3 bins, 2 bins to store the items of the first column and another bin to store the rightmost $\Gamma - 1$ columns. Although in this example all items $i \in [n]$ are set to have ratios $\hat{a}_i/\bar{a}_i = c$, the example can be extended in a straightforward manner with slight perturbations of the item sizes so that the ratios will be strictly decreasing for the items ordered from left-to-right and top-to-bottom in this matrix. □

4 First ideas to get a constant ratio for ΓRBP

We maintain the assumption that the items are ordered according to (2).

4.1 Attempts to get a constant ratio

We discuss below some natural arguments to get constant ratios.

Attempt 1: using a classical size argument NEXT-FIT without a particular ordering applied to instance of Section 3 leads to a solution with $k' = n/2$ bins (at the end of Step 1) where $f_\Gamma(b_j) > 1$ for each bin, while $\mathsf{OPT} = 2$. This example shows that even if all bins are "full" (relatively to f_Γ), it does not provide a lower bound on the optimal number of bins. Moreover, as shown in Theorem 4, none of the two orders considered in the previous section leads to a constant ratio using NEXT-FIT.

Attempt 2: using the duality with makespan minimization Given input $\mathcal{I}$, we could guess $k^* = \mathsf{OPT}(\mathcal{I})$, and then consider the input $(\mathcal{I}, k^*)$ as an input of robust makespan minimzation (which was studied in [6]). Using any ρ-approximation for the later problem (for example $\rho = 3$ in [6]), we could get in polynomial time a solution with k^* bins an such that $f_\Gamma(b_j) \leq \rho$. The last step would be to convert this solution into a solution of ΓRBP by unpacking each bin (with $f_\Gamma(b_j) \leq 3$) into several bins b'_j with $f_\Gamma(b'_j) \leq 1$. However, even if ρ were arbitrarily close to 1, it is not possible to bound (for a fixed j) the number of bins b'_j by a constant as showed in the instance containing n items $(\frac{\epsilon}{n}, 1 - \frac{\epsilon}{n})$ and $\Gamma = 1$. While all items fit into a single bin with capacity lower than $1 + \epsilon$, they require n bins of capacity 1 to be packed.

Attempt 3: guessing the profile of an optimal solution Let $\mathcal{I}$ be a input of ΓRBP. Given a solution $s = \{b_j, j \in [k]\}$ for this input, we define $P(s) = \{\Gamma(b_j), j \in [k]\}$ as the profile of s and $\tilde{P}_j(s) = \{i \mid i \in \Gamma(b_\ell) \text{ for some } \ell \in [j]\}$ as all deviating items in the first j bins. Let $s^* = \{b^*_j, j \in [k^*]\}$ be an optimal solution. To get some insight on the problem, let us assume that we know $P(s^*)$ (even if this cannot be guessed in polynomial time). We show how we can use $P(s^*)$ to get a 2-approximation algorithm. Without loss of generality, we can always assume that $|\Gamma(b^*_j)| = \Gamma$ for any j, as otherwise we can add $\Gamma - |\Gamma(b^*_j)|$ dummy items of size $(0, 0)$ to b^*_j. Remember that the items are sorted in non-increasing order of their deviating values ($\hat{a}_i \geq \hat{a}_{i+1}$). For any $j \in [k^*]$, let $i^*_j = \max(\Gamma(b^*_j))$ be the smallest (in term of $\hat{a}$ value) deviating item of bin j (when $\Gamma = 1$, $\{i^*_j\} = \Gamma(b^*_j)$ as in the previous section). Without loss of generality, let us assume that bins are sorted such that $i^*_j \geq i^*_{j+1}$. Now, given $P(s^*)$, in the first phase we construct a solution s by packing items of $P(s^*)$ as they were packed in s^*, meaning that we define $b_j = \Gamma(b^*_j)$ for $j \in [k^*]$. Let $X = [n] \setminus \bigcup_{j \in [k^*]} \Gamma(b^*_j)$ be the set of remaining items. We now pack X in the following second phase, starting with $j = 1$. Notice that in the description of the algorithm below, we consider that for $j \in [k^*]$, b_j already contains $\Gamma(b^*_j)$, whereas for any $j > k^*$, b_j is initially empty.

Step 1 pack items of X (by decreasing $\hat{a}$ values) in b_j until $f_\Gamma(b_j) > 1$ or $X = \emptyset$
Step 2 if $X \neq \emptyset$, $j = j + 1$, and go to step 1.

Let j be the bin such that X is empty after filling b_j. Let k' be the number of bins used by this algorithm. Notice that if $j \leq k^*$ then $k' = k^*$ (because of the pre-packing of item of $P(s^*)$), and otherwise $k' = j$.

Lemma 6. *$k' \leq k^*$, implying a 2-approximation as we can convert the solution of* Next-Fit *into a feasible solution of $2k'$ bins by repacking the last added item in each bin in a separate bin.*

Proof. Assume by contradiction that $k' > k^*$. As an item $i \in [n] \setminus \tilde{P}_{k^*}(s^*)$ does not deviate in s^*, we need to ensure that this is also the case in s. Let us prove by induction on j that the items packed greedily in Step 1 satisfy

$$\hat{a}_i \leq \hat{a}_{i_j^*}, \forall i \in b_j \setminus \tilde{P}_{k^*}(s^*), j \in [k^*]. \tag{14}$$

Let $j = 1$, and suppose there is $i \in b_1 \setminus \tilde{P}_{k^*}(s^*)$ such that $\hat{a}_i > \hat{a}_{i_1^*}$. Then, because $\hat{a}_{i_1^*} \geq \hat{a}_{i_j^*}$ for $j > 1$, $\hat{a}_i > \hat{a}_{i_j^*}$ for each j so $i \in \tilde{P}_{k^*}(s^*)$, a contradiction. Now, consider bin b_{j+1}. By induction, we have that $\sum_{\ell \in [j]} \overline{a}(b_\ell) + \hat{a}(\tilde{P}_j(s)) > j \geq \sum_{\ell \in [j]} \overline{a}(b_\ell^*) + \hat{a}(\tilde{P}_j(s^*))$, so $\tilde{P}_j(s) = \tilde{P}_j(s^*)$ implies

$$\sum_{\ell \in [j]} \overline{a}(b_\ell) > \sum_{\ell \in [j]} \overline{a}(b_\ell^*). \tag{15}$$

Let X_j be the set of items of X left after packing bin b_j by the above procedure and X_j^* be the the set of items of X left after the optimal solution packs bin b_j^*. Inequality (15) and the ordering used in Step 1 imply that $\lambda \geq \lambda^*$, where $\lambda = \min(X_j)$ and $\lambda^* = \min(X_j^*)$. Therefore, if there exists $i \in b_{j+1} \setminus \tilde{P}_{k^*}(s^*)$ such that $\hat{a}_i > \hat{a}_{i_{j+1}^*}$, then $\hat{a}_{\lambda^*} \geq \hat{a}_\lambda \geq \hat{a}_i > \hat{a}_{i_{j+1}^*}$, and thus $\hat{a}_{\lambda^*} > \hat{a}_{i_\ell^*}$ for any $l \in [\![j+1, k^*]\!]$, which is a contradiction as item λ^* is in X and thus does not deviate in the considered optimal.

Now that Property (14) is proved, let us get our contradiction from $k' > k^*$. Indeed, if $k' > k^*$ then $\sum_{i \in [n]} \overline{a}_i > k^* - \hat{a}(\tilde{P}_{k^*}(s^*)) \geq \sum_{j \in [k^*]} \overline{a}(b_j^*)$ where the first inequality follows from $f_\Gamma(b_j) > 1$ for $j \in [k^*]$ and Property (14), and the second one follows from $\sum_{j \in [k^*]} \overline{a}(b_j^*) + \hat{a}(\tilde{P}_{k^*}(s^*)) \leq k^*$. This implies a contradiction. □

Even if the above procedure relies on a guessing step which is not polynomial, its core idea has similarities with both the analysis of Next-Fit in the proof of Theorem 2 and with the DP algorithm detailed later in this paper, where we only guess the deviating item with the smallest deviation of each bin (one at a time), and we pack $\Gamma - 1$ items "better" than the one packed in $P(s^*)$, at the expense of a few extra bins.

4.2 Restricting our attention to small items.

We define ΓRBP with small values as the ΓRBP problem restricted to inputs where for any $i \in [n]$, $\hat{a}_i \leq \frac{1}{\Gamma}$ and $\hat{a}_i \leq \frac{1}{\Gamma}$. Below we give a justification for restricting our attention to ΓRBP with small values.

Lemma 7. *Any polynomial ρ-approximation for ΓRBP with small values implies a polynomial $(\rho + \rho_{bp})$-approximation for ΓRBP, where ρ_{bp} is the best known ratio of a polynomial time approximation for classical bin-packing.*[4]

Proof. Given an instance $\mathcal{I}$ of ΓRBP, we define the small items $\mathcal{S} = \{i \in [n] : \overline{a}_i \leq 1/\Gamma \text{ and } \hat{a}_i \leq 1/\Gamma\}$ and the large item as $\mathcal{B} = [n] \setminus \mathcal{S}$. We use our ρ-approximation algorithm to pack $\mathcal{S}$ into $k_{\mathcal{S}}$ bins, implying $k_{\mathcal{S}} \leq \rho\mathsf{OPT}(\mathcal{S}) \leq \rho\mathsf{OPT}(\mathcal{I})$. Then, we observe that in any packing of $\mathcal{B}$, each bin contains no more than Γ items, so that all items deviate in these bins. Hence, ΓRBP for instance $(\mathcal{B}, \Gamma)$ is equivalent to the classical bin-packing problem for items $\mathcal{B}'$ where the weight of each item $i \in \mathcal{B}'$ is given by $\overline{a}_i + \hat{a}_i$. This implies that $\mathsf{OPT}_{bp}(\mathcal{B}') = \mathsf{OPT}(\mathcal{B})$ (where OPT_{bp} denotes the optimal value in classical bin-packing), and that any solution for $\mathcal{B}'$ is a solution of $\mathcal{B}$. Thus, we use a ρ_{bp}-approximation algorithm for classical bin-packing to pack $\mathcal{B}'$ in $k_{\mathcal{B}}$ bins, and use the same packing for items in $\mathcal{B}$. Note that $k_{\mathcal{B}} \leq \rho_{bp}\mathsf{OPT}_{bp}(\mathcal{B}) = \rho_{bp}\mathsf{OPT}(\mathcal{B}) \leq \rho_{bp}\mathsf{OPT}(\mathcal{I})$. We obtain a packing of $\mathcal{I}$ with cost $k_{\mathcal{S}} + k_{\mathcal{B}} \leq (\rho + \rho_{bp})\mathsf{OPT}(\mathcal{I})$. □

Observation 1 *Given an instance $\mathcal{I}$ to the ΓRBP with small values, any subset $X \subseteq [n]$ can be packed in $\lceil \frac{|X|}{\Gamma/2} \rceil$ bins.*

Notice that instances with small items are not easier to approximate by Next-Fit because instance (13) from Section 3 uses small items.

4.3 Guessing of the full profile and considering only small items.

Let us now explain why mixing the two previous ideas is promising. As in attempt 3 where we know the full profile, we want to construct for any j bins $\{b_1, \ldots, b_j\}$ such that their total $\overline{a}$ is larger than the total value of $\overline{a}$ packed by the first j bins of s^* (the considered optimal solution), as in inequality (15). Instead of guessing the full profile $P(s^*)$, we want to design a DP algorithm (that guesses i_{j^*} one at the time) with the following flavor. Start with $j = 1$.

- guess item i^*_j, the smallest (in $\hat{a}$ value) deviating item of b^*_j, and pack it in b_j
- then, as we do not know which are the $\Gamma - 1$ other deviating items in b^*_j and we want to pack more $\overline{a}$ area, packs separately $\Gamma - 1$ items with larger $\overline{a}$ values (among items with $\hat{a}$ values greater than $\hat{a}_{i^*_j}$). Consider that these $\Gamma - 1$ items are put in the "trash", and at the very end we will pack all items of the trash in a few bins
- keep filling bin b_j greedily (by non increasing $\hat{a}$ values) until exceeding 1
- make a recursive call with $j + 1$

[4] In general, if we have a polynomial time additive approximation algorithm using $OPT + f(OPT)$ bins and polynomial time ρ-approximation algorithm for ΓRBP with small values then our algorithm uses $OPT(\rho + 1) + f(OPT)$ bins for ΓRBP in polynomial time.

If s^* uses k^* bins, we wish to output a solution s with k^* bins exceeding one, and $(\Gamma-1)k^*$ items in the trash. This almost feasible solution can be converted into a regular one with $3k^*$ bins by removing one item from each bin and adding them to the trash, and packing the Γk^* items of the trash into $2k^*$ bins, which is possible according to Observation 1. This sketches the core ideas of the DP. However, the actual DP presented below needs to be more involved the two following reasons. Consider $j = 1$ to simplify and let $B = [\![1, i_1^* - 1]\!]$.

Firstly, notice that in s^*, items of B could be packed (as deviating items) in another bin than b_1^*, and we may have $|B| > \Gamma - 1$. Thus, instead of trashing only $\Gamma - 1$ items of B, we have to trash all of them, and count the number of trashed items to ensure that at the end at most $(\Gamma-1)k^*$ items are trashed. To summarize, the trash will represent the union of the $(\Gamma - 1)$ larger (in $\hat{a}$ values) deviating items of each bin. Moreover, we want to maintain that the $\overline{a}$ area of trashed items in s is larger than the $\overline{a}$ area of deviating items in s^*.

Secondly, notice that in s^*, items of $[\![i_1^* + 1, i_2^* - 1]\!]$ are either in b_1^* as non-deviating items or in a b_j^*, $j \geq 2$ as deviating items (meaning that they are trashed items in s). Thus, if we incorrectly pack some of these items in b_1 instead of trashing them, these items will not be available when considering b_2, and we may not be able to ensure then that trashed items in s have a larger $\overline{a}$ value than the deviating items in s^*.

We describe in the next section the correct version of the DP. To that end, we first need to introduce formally the notion of trash.

5 Approximating ΓRBP with small values

Bin-packing with trash For any $X \subseteq [n]$, we define $\tilde{a}_\Gamma(X) = \Gamma \hat{a}_1(X)$ ($\tilde{a}_\Gamma(X)$ is Γ times the largest deviating value of an item in X) and $\tilde{f}(X) = \overline{a}(X) + \tilde{a}_\Gamma(X)$. We introduce next a decision problem ΓRBP-T related to ΓRBP.

> ΓRBP-T (Robust bin-packing with trash)
> *Input:* $(\mathcal{I}, k, t)$ where $\mathcal{I}$ is an input of ΓRBP (where each item $(\overline{a}_i, \hat{a}_i)$ satisfies $\hat{a}_i \leq 1/\Gamma$ and $\overline{a}_i \leq 1/\Gamma$), and k, t are two integers.
> *Output:* Decide if a solution exists, where a solution is a partition of the set of items into $k+1$ sets $b_1, \ldots, b_k$ and T (called the trash) such that:
> - $\tilde{f}(b_j) \leq 1$ for each $j = 1, \ldots, k$
> - $|T| \leq t$

Notice that although each item is small in ΓRBP-T, it is possible to have an item i such that $\tilde{f}(\{i\}) > 1$, implying that i must be put in the trash. We show below how deciding ΓRBP-T is enough to approximate ΓRBP.

Lemma 8. *For any input $\mathcal{I}$ of ΓRBP and $k^* = \mathsf{OPT}(\mathcal{I})$, $(\mathcal{I}, k^*, (\Gamma-1)k^*)$ is a yes input of ΓRBP-T.*

Proof. Given an optimal solution of size k^* of ΓRBP problem we create a solution to ΓRBP-T problem as follows. Let b_j^* be a bin of the considered optimum. Let N_j be the non-deviating items of b_j^*, e.g. $b_j^* = N_j \cup \Gamma(b_j^*)$. Let

$X = \max(\Gamma(b_j^*))$ (the smallest deviating item of b_j^*) if $|\Gamma(b_j^*)| = \Gamma$ and $X = \emptyset$ otherwise. We define $b'_j = N_j \cup X$, and add items of $Y = b_j^* \setminus b'_j$ into the trash. Notice Y is either the set of $\Gamma - 1$ largest deviating object of b_j^*, or is equal to $\Gamma(b_j^*)$ when $|\Gamma(b_j^*)| < \Gamma$.This is a feasible solution for ΓRBP-T problem as $\tilde{f}(b'_j) = \overline{a}(b'_j) + \tilde{a}_\Gamma(b'_j)$, where $\overline{a}(b'_j) \leq \overline{a}(b_j^*)$ and $\tilde{a}_\Gamma(b'_j) = \tilde{a}_\Gamma(X) \leq \hat{a}_\Gamma(b_j^*)$, and as there are at most $(\Gamma - 1)k^*$ items in the trash. □

Lemma 9. *For any input $\mathcal{I}$ of Γ*RBP *and integer k, given a solution of $(\mathcal{I}, k, \Gamma k)$ of Γ*RBP-T*, we can compute in polynomial time a solution of $3k$ bins for $\mathcal{I}$.*

Proof. Given a solution $b_1, \ldots, b_k, T$ for $(\mathcal{I}, k, \Gamma k)$ of ΓRBP-T the bins remain feasible in ΓRBP as $f_\Gamma(b_j) = \overline{a}(b_j) + \hat{a}_\Gamma(b_j) \leq \overline{a}(b_j) + \tilde{a}(b_j) = \tilde{f}(b_j)$. Then, Observation 1 implies that the trash T can be packed into $\lceil k\Gamma/(\Gamma/2) \rceil \leq 2k$ additional bins. □

A DP algorithm for ΓRBP-T The objective of this section is to define a DP algorithm that will be used to decide the ΓRBP-T problem. To this aim, we define G-ΓRBP-T (generalized robust bin-packing with trash), an optimization problem that the DP algorithm will solve in a relaxed way. To define G-ΓRBP-T, we fix a sorted instance $\mathcal{I}$ of ΓRBP according to (2) and an integer k.

G-ΓRBP-T (generalized robust bin-packing with trash)
Input: $\mathcal{I} = (q, t, \ell)$, where $q \in [n]_0$, $t \in [(\Gamma - 1)k]_0$, and $\ell \in [k + 1]$.
Output: A feasible solution s is a partition of $[\![q, n]\!]$ into $k - \ell + 3$ sets (b_j for $j \in [\![\ell, k]\!]$, b_0 and T), such that
- for any $j \in [\![\ell, k]\!]$, $\tilde{f}(b_j) \leq 1$ (the $k - \ell + 1$ regular bins must respect the constraint of ΓRBP-T)
- $|T| \leq t$ (we only allow t items in the trash)
- $\min(b_\ell) = q$ (meaning that the deviating item of b_ℓ is q as items are sorted in non increasing order of $\hat{a}$ values)

Minimize: $c(s) = \overline{a}(b_0)$ (in bin b_0 we only count $\overline{a}$ values)

The objective of G-ΓRBP-T is to pack a part (defined by $[\![q, k]\!]$) of an ΓRBP-T instance given a fixed budget of resources (the number of bins and the size of the trash) while minimizing the total nominal size of items in the dummy bin b_0. The last constraint (the deviating item of b_ℓ is q) may appears artificial at first sight, but comes from the fact that the DP will guess at each new bin the largest items that should be packed in it, and thus this constraints ensure that any optimal solution must pack q in b_ℓ as well.

Definition 1 (almost feasible solution). *We say that a bin b **exceeds by at most one item** iff $\tilde{f}(b) > 1$ and $\tilde{f}(b \setminus \{i\}) \leq 1$ where $i = \max\{b\}$. Given an input (q, t, ℓ) of* G-ΓRBP-T*, we say that a solution is **almost feasible** iff all the above constraints of* G-ΓRBP-T *are respected, except that for any $j \in [\![\ell, k]\!]$, we allow that b_j exceeds by at most one item instead of $\tilde{f}(b_j) \leq 1$.*

The relation between G-ΓRBP-T and ΓRBP-T is characterized in the following two lemmas.

Lemma 10. *For any $\mathcal{I}$ input of ΓRBP and k such that $(\mathcal{I}, k, (\Gamma-1)k)$ is a yes input of ΓRBP-T, there exists q and t such that $\mathsf{OPT}(q,t,1) = 0$.*

Proof. Consider a solution of $(\mathcal{I}, k, (\Gamma-1)k)$ (of ΓRBP-T). Let $q = \min\left(\bigcup_{j=1}^{k} b_j\right)$ (the first item starting from 1 which is placed into a bin, say bin b_1 without loss of generality). Let $t = (\Gamma-1)k - (q-1)$. By definition of q, items in $[q-1]$ are placed in T, and thus it means that $\{b_1, \ldots, b_k\}$ and $T \setminus [q-1]$ is a solution of $(q,t,1)$ of cost 0. □

Lemma 11. *Let us fix $\mathcal{I}$ an input of ΓRBP and k an integer. For any $q \in [n], t \in [(\Gamma-1)k]$, given an almost feasible solution of $\mathcal{I}' = (q,t,1)$ of cost 0 for G-ΓRBP-T, we can compute in polynomial time a solution of $(\mathcal{I}, k, \Gamma k)$ of ΓRBP-T.*

For any $q \in [(\Gamma-1)k], t = (\Gamma-1)k-(q-1)$, given an almost feasible solution of $\mathcal{I}' = (q,t,1)$ of cost 0 for G-ΓRBP-T, we can compute in polynomial time a solution of $(\mathcal{I}, k, \Gamma k)$ of ΓRBP-T.

Proof. Let us consider an almost feasible solution of G-ΓRBP-T for $\mathcal{I}'$ of cost 0. We remove the item of larger index of each bin, and add it to T. We also add items $[q-1]$ in T. Let b'_j and T' be respectively the new bins and set T obtained after performing this operation. We now have $\tilde{f}(b'_j) \leq 1$ for any $j \in [k]$ (as b_j exceeds by at most one item), and $|T'| \leq \Gamma k$, concluding the proof. □

Thus, Lemmas 9 and 11 show that providing an almost feasible solution for $(q,t,1)$ of cost 0 for G-ΓRBP-T implies a solution of size $3k$ for ΓRBP.

Let us now define a DP algorithm $DP(\mathcal{I})$ ($\mathcal{I}$ is an input of G-ΓRBP-T) that provides an **almost feasible** solution s with $c(s) \leq \mathsf{OPT}(\mathcal{I})$ (where $\mathsf{OPT}(\mathcal{I})$ is by definition the optimal cost of a **feasible** solution). We provide below a gentle description of the DP. Given an instance $\mathcal{I} = (q,t,\ell)$, the DP starts by guessing (q^*, t^*), where

- $q^* = \min(b^*_{l'})$ for a bin $b^*_{l'}$ with $l' \in [\![l+1, k]\!]$ of an optimal solution s^*
- t^* is the number of items trashed from X^* in s^*, where $X^* = [\![q, q^*-1]\!]$
- Notice that in s^* items of X^* must by placed in b^*_l, b^*_0 or T^*. We mimic the optimal in the current call of the DP by packing X^* in b_l, b_0 and T.
- To that end, the DP:
  - packs q to b_l (as required by the corresponding constraint of G-ΓRBP-T),
  - packs the t^* largest (in terms of $\overline{a}$) remaining items of X^* to the trash
  - packs the remaining items of X^* into b_ℓ until $\tilde{f}(b_\ell) > 1$ or $X^* = \emptyset$
  - packs the remaining items of X^* into b_0 until $X^* = \emptyset$

We discuss next where the others items (of $[\![q^*, n]\!]$) are packed. Notice that in s^*, bin b^*_l may contain items of $[\![q^*, n]\!]$, and thus the DP may also have to pack items of $[\![q^*, n]\!]$ into b_l. The key is that the decision of which items of $[\![q^*, n]\!]$ to pack into b_l is not taken at this step of the algorithm but only later (to avoid packing

in b_l items of large $\overline{a}$ value that are in the trash in s^*). To allow this decision to be taken later, let Δ_b be the size of the empty space in b_l after packing X^* as described above, and let $b_0^{X^*} = b_0 \cap X^*$. After the previous steps, the DP makes a recursive call to get a solution $\tilde{s}$ that packs $[\![q^*, n]\!]$ into regular bins, a trash, and a dummy bin $\tilde{b}_0$. So far solution $\tilde{s}$ has not benefited from the empty space Δ_b. However, we can unpack items from $\tilde{b}_0$ to b_ℓ while ensuring that these items do not deviate in b_ℓ (as all these items have index greater than q).

5.1 The DP solving G-ΓRBP-T

Let us first define formally the DP in Algorithm 2. We refer the read to Figure 1 where the behavior of DP is depicted.

1 **if** $q = n + 1$ **then return** $\emptyset$
2 Initialize s to any almost feasible solution.
3 **for** $g = (q', t') \in [\![q + 1, n]\!] \times [t]_0$ **do**
4 Let $X' = [\![q, q' - 1]\!]$
5 Pack q to b_ℓ
6 Add T' to T, where $T' = \{t'$largest items (in terms of $\overline{a}$) of $X' \setminus \{q\}\}$
`// breaking ties arbitrarily`
7 Pack $X' \setminus (\{q\} \cup T')$ to b_ℓ until $\tilde{f}(b_\ell) > 1$ or $X' \setminus (\{q\} \cup T') = \emptyset$
8 Let X'_0 be the remaining items `// $X'_0$ may be empty`
9 Pack X'_0 to b_0
10 Let $\mathcal{I}(g) = (q', t - t', \ell + 1)$
11 Call $\mathbf{DP}(\mathcal{I}(g))$, outputting $\tilde{s} = \{\tilde{b}_0\} \cup \{\tilde{b}_j, j \in [\![\ell + 1, k]\!]\} \cup \tilde{T}$.
12 Unpack b'_0 to b_ℓ until $\tilde{f}(b_\ell) > 1$ or all items of b'_0 are unpacked, then unpack the potentially remaining items of b'_0 to b_0.
13 Let s_g the obtained solution
14 **if** $c(s_g) < c(s)$ **then** $s = s_g$
return: s

Algorithm 2: $\mathbf{DP}(q, t, \ell)$

Let us introduce notations to describe the packing of the DP. Let $b_\ell^{X'} = b_\ell \cap X'$ be the items of X' packed in b_ℓ by the DP, and let $\Delta_b = 1 - \tilde{f}(b_\ell^{X'})$ (Δ_b could be negative). In the same way, let $b_0^{X'} = b_0 \cap X'$. Let $c_0 = \overline{a}(b_0^{X'})$. Notice that c_0 corresponds to the total non deviating size of items packed in b_0 after step 9 of Algorithm 2. Let us now prove that the specification of the algorithm is true.

Lemma 12. *For any $\mathcal{I}$ input of* G-ΓRBP-T*, $DP(\mathcal{I})$ provides an almost feasible solution of cost at most $\mathsf{OPT}(\mathcal{I})$.*

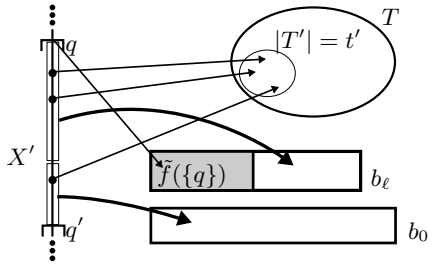

Fig. 1. DP handling guess (q', t'), starting from item q.

Proof. The proof is by induction on ℓ. Suppose now that the claim is true for $\ell + 1$. Let s^* be an optimal (feasible) solution to $\mathcal{I}$ with bins b_j^* and trash T^*. Let $q^* = \min\left(\bigcup_{\ell' \in [\![\ell+1,k]\!]} b^*_{\ell'}\right)$ be the first item packed in a regular bin different from ℓ, and let $X^* = [\![q, q^* - 1]\!]$. Notice q is necessarily packed (by s^*) in b_ℓ^* as it is a constraint of the problem, and other items of X^* may only be packed in b_ℓ^*, b_0^* or T^*. Let $T^{*X^*} = T^* \cap X^*$, $b_\ell^{*X^*} = b_\ell^* \cap X^*$, and $b_0^{*X^*} = b_0^* \cap X^*$. Define $c_0^* = \overline{a}(b_0^{*X^*})$, $\Delta_b^* = 1 - \tilde{f}(b_\ell^{*X^*})$.

Let $g^* = (q^*, t - |T^{*X^*}|)$. Recall that $\mathcal{I}(g^*) = (q^*, t - |T^{*X^*}|, \ell + 1)$. From the existence of s^* we will deduce the existence of the following solution for $\mathcal{I}(g^*)$. Let $\tilde{s^*} = \{\tilde{T^*}, \tilde{b_0^*}\} \bigcup \{\tilde{b_j^*}, j \in [\![\ell + 1, k]\!]\}$ be the solution obtained from s^* by

- keeping bins $\ell + 1$ to k unchanged ($\tilde{b_j^*} = b_j^*$ for $j \in [\![\ell + 1, k]\!]$)
- removing items of X^* in T^* ($\tilde{T^*} = T^* \setminus X^*$)
- removing items of X^* from b_0^* and adding to b_0^* items of $b_\ell^* \setminus X^*$ ($\tilde{b_0^*} = (b_0^* \setminus X^*) \cup (b_\ell^* \setminus X^*)$)

Observe that $\tilde{s^*}$ is a feasible solution of $\mathcal{I}(g^*)$. Moreover, as $b_0^* = (b_0^{*X^*} \cup \tilde{b_0^*}) \setminus (b_\ell^* \setminus X^*)$ and as $\overline{a}(b_\ell^* \setminus X^*) \leq \Delta_b^*$, we get that $c(s^*) = \overline{a}(b_0^*) \geq c_0^* + c(\tilde{s^*}) - \Delta_b^* \geq c_0^* + \mathsf{OPT}(\mathcal{I}(g^*)) - \Delta_b^*$.

On the other hand, as the algorithm tries all possible guesses, we have $c(s) \leq c(s_{g^*})$. Notice that for guess g^*, the set X' defined in the algorithm is equal to X^*. Observe that $c(s_{g^*}) \leq c_0 + c(\tilde{s}) - \Delta_b$. Indeed, suppose first that $\Delta_b \geq 0$, implying that $c_0 = 0$. When we unpack items from $\tilde{b}_0$ to b_ℓ, as these items will not deviate in b_ℓ and as we add items until $\tilde{f}(b_\ell) > 1$, we will either add the whole set $\tilde{b}_0$ (implying $c(s_{g^*}) = 0$), or move a volume (in $\overline{a}$) of at least Δ_b (implying $c(s_{g^*}) \leq c(\tilde{s}) - \Delta_b$). In the other case ($\Delta_b \leq 0$), we get $c(s_{g^*}) = c_0 + c(\tilde{s})$ as no item will be unpacked, implying also the desired expression. Moreover, by the induction hypothesis, we have $c(\tilde{s}) \leq \mathsf{OPT}(\mathcal{I}(g^*))$.

Now, to prove that $c(s_{g^*}) \leq c(s^*)$, it remains to prove that $c_0 - \Delta_b \leq c_0^* - \Delta_b^*$. We have

$$
\begin{aligned}
c_0 - \Delta_b \leq c_0^* - \Delta_b^* &\Leftrightarrow c_0 + \tilde{f}(b_\ell^{X^*}) \leq c_0^* + \tilde{f}(b_\ell^{*X^*}) \\
&\Leftrightarrow c_0 + \overline{a}(b_\ell^{X^*}) \leq c_0^* + \overline{a}(b_\ell^{*X^*}) \text{ as the deviating item is } q \\
&\qquad \text{in both optimal and algorithm bins} \\
&\Leftrightarrow \overline{a}(X^* \setminus T') \leq \overline{a}(X^* \setminus T^*) \text{ as } X^* \setminus (\{f\} \cup T') = b_0^{X^*} \cup b_\ell^{X^*} \\
&\qquad \text{and } X^* \setminus (\{f\} \cup T^*) = b_0^{*X^*} \cup b_\ell^{*X^*}
\end{aligned}
$$

The last inequality is true as $|T'| = |T^*|$ and T' contains the largest (in $\overline{a}$) items, implying the desired result. □

Lemma 13. *There is a* 3*-approximation for* ΓRBP *with small values running in* $\mathcal{O}(n^6 log(n))$.

Proof. Given an input $\mathcal{I}$ of ΓRBP, we perform a binary search on k. Given a k, if there exists q and t such that $c(DP(f,t,1)) \leq 0$ then we decrease k, otherwise we increase it. Let us now prove that this is a 3-approximation algorithm. Let $k^* = \mathsf{OPT}(\mathcal{I})$ be the optimal value of ΓRBP. According to Lemmas 8 and 9, this implies that $(\mathcal{I}, k^*, (\Gamma - 1)k^*)$ is a YES-instance of ΓRBP-T, implying in turn that there exists f, t such that $\mathsf{OPT}(f,t,1) = 0$. Then, according to Lemma 12, $DP(f,t,1)$ outputs s, an almost feasible solution with $c(s) \leq \mathsf{OPT}(f,t,1) = 0$. According to Lemma 10 and 11, we can compute a solution of $(\mathcal{I}, k^*, \Gamma k^*)$ for ΓRBP-T, and then a solution with $3k^*$ bins for ΓRBP.

Let us now turn to the running time. The different number of inputs of the DP is $\mathcal{O}(n^3)$, and the running time for a fixed input is $\mathcal{O}(n^3)$. Thus, taking into account the binary search, we get a running time in $\mathcal{O}(n^6 log(n))$. □

According to Lemma 7, the following Theorem is now immediate using a $\frac{3}{2}$-approximation for classical bin-packing (see for example in [15]) as a black box.

Theorem 5. *There is a* 4.5*-approximation for* ΓRBP *running in* $\mathcal{O}(n^6 log(n))$.

Bibliography

[1] Aissi, H., Bazgan, C., Vanderpooten, D.: Min–max and min–max regret versions of combinatorial optimization problems: A survey. European journal of operational research **197**(2), 427–438 (2009)

[2] Álvarez-Miranda, E., Ljubic, I., Toth, P.: A note on the bertsimas & sim algorithm for robust combinatorial optimization problems. 4OR **11**(4), 349–360 (2013). https://doi.org/10.1007/s10288-013-0231-6, https://doi.org/10.1007/s10288-013-0231-6

[3] Balogh, J., Békési, J., Dósa, G., Sgall, J., van Stee, R.: The optimal absolute ratio for online bin packing. In: Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (2015)

[4] Balogh, J., Békési, J., Dósa, G., Epstein, L., Levin, A.: A New and Improved Algorithm for Online Bin Packing. In: Azar, Y., Bast, H., Herman, G. (eds.) ESA. LIPIcs, vol. 112, pp. 5:1–5:14. Dagstuhl, Germany (2018)

[5] Bertsimas, D., Sim, M.: Robust discrete optimization and network flows. Math. Program. **98**(1-3), 49–71 (2003)

[6] Bougeret, M., Pessoa, A.A., Poss, M.: Robust scheduling with budgeted uncertainty. Discrete Applied Mathematics (2018), to appear

[7] Dexter, F., Macario, A., Traub, R.D.: Which algorithm for scheduling add-on elective cases maximizes operating room utilization? use of bin packing algorithms and fuzzy constraints in operating room management. Anesthesiology **91**, 1491–1500 (Nov 1999)

[8] Goetzmann, K., Stiller, S., Telha, C.: Optimization over integers with robustness in cost and few constraints. In: WAOA. pp. 89–101 (2011)

[9] Gounaris, C.E., Wiesemann, W., Floudas, C.A.: The robust capacitated vehicle routing problem under demand uncertainty. Operations Research **61**(3), 677–693 (2013)

[10] Johnson, D.S., Demers, A., Ullman, J.D., Garey, M.R., Graham, R.L.: Worst-case performance bounds for simple one-dimensional packing algorithms. SIAM Journal on Computing **3**(4), 299–325 (1974)

[11] Karmarkar, N., Karp, R.M.: An efficient approximation scheme for the one-dimensional bin-packing problem. In: Proc. 23rd Annual Symp. Foundations of Computer Science (sfcs 1982). pp. 312–320 (Nov 1982)

[12] Kasperski, A., Zieliński, P.: On the approximability of minmax (regret) network optimization problems. Information Processing Letters **109**(5), 262–266 (2009)

[13] Kasperski, A., Zielinski, P.: On the approximability of robust spanning tree problems. Theor. Comput. Sci. **412**(4-5), 365–374 (2011). https://doi.org/10.1016/j.tcs.2010.10.006, https://doi.org/10.1016/j.tcs.2010.10.006

[14] Poss, M.: Robust combinatorial optimization with knapsack uncertainty. Discrete Optimization **27**, 88–

102 (2018). https://doi.org/10.1016/j.disopt.2017.09.004, https://doi.org/10.1016/j.disopt.2017.09.004

[15] Simchi-Levi, D.: New worst-case results for the bin-packing problem **41**, 579–585 (1994)

[16] Song, G., Kowalczyk, D., Leus, R.: The robust machine availability problem–bin packing under uncertainty. IISE Trans. pp. 1–35 (2018)

[17] Tadayon, B., Smith, J.C.: Algorithms and complexity analysis for robust single-machine scheduling problems. J. Scheduling **18**(6), 575–592 (2015)