arXiv:2205.12828v3 [cs.DS] 28 Apr 2023

Improved Approximations for Vector Bin Packing via Iterative Randomized Rounding

Ariel Kulik* Matthias Mnich† Hadas Shachnai‡

Abstract

We study the d-DIMENSIONAL VECTOR BIN PACKING (dVBP) problem, a generalization of BIN PACKING with central applications in resource allocation and scheduling. In dVBP, we are given a set of items, each of which is characterized by a d-dimensional *volume* vector; the objective is to partition the items into a minimum number of subsets (bins), such that the total volume of items in each subset is at most 1 in each dimension.

Our main result is an asymptotic approximation algorithm for dVBP that yields a ratio of $(1+\ln d-\chi(d)+\varepsilon)$ for all $d \in \mathbb{N}$ and any $\varepsilon > 0$; here, $\chi(d)$ is some strictly positive function. This improves upon the best known asymptotic ratio of $(1 + \ln d + \varepsilon)$ due to Bansal, Caprara and Sviridenko (SICOMP 2010) for any $d > 3$. By slightly modifying our algorithm to include an initial matching phase and applying a tighter analysis we obtain an asymptotic approximation ratio of $\left(\frac{4}{3} + \varepsilon\right)$ for the special case of $d = 2$, thus substantially improving the previous best ratio of $\left(\frac{3}{2} + \varepsilon\right)$ due to Bansal, Eliáš and Khan (SODA 2016).

Our algorithm iteratively solves a configuration LP relaxation for the residual instance (from previous iterations) and samples a small number of configurations based on the solution for the configuration LP. While iterative rounding was already used by Karmarkar and Karp (FOCS 1982) to establish their celebrated result for classic (one-dimensional) BIN PACKING, iterative *randomized* rounding is used here for the first time in the context of (VECTOR) BIN PACKING. Our results show that iterative randomized rounding is a powerful tool for approximating dVBP, leading to simple algorithms with improved approximation guarantees.

*CISPA Helmholtz Center for Information Security, Germany. ariel.kulik@cispa.de

†Hamburg University of Technology, Institute for Algorithms and Complexity, Hamburg, Germany. matthias.mnich@tuhh.de

‡Computer Science Department, Technion, Haifa, Israel. hadas@cs.technion.ac.il

Contents

1 Introduction **1**
- 1.1 Our Contribution . . . 1
- 1.2 Related Work . . . 2
- 1.3 The Algorithm . . . 3
- 1.4 Improved Algorithm for 2VBP . . . 6
- 1.5 Technical Contribution . . . 8
- 1.6 Organization . . . 9

2 Preliminaries **9**
- 2.1 Probability Space . . . 9
- 2.2 McDiarmid's Concentration Bound . . . 10
- 2.3 First-Fit . . . 11

3 Improved Asymptotic Approximation for dVBP **12**
- 3.1 Proof of Theorem 3.2 . . . 13
- 3.2 The Weak Structural Property . . . 24

4 Asymptotic $\left(\frac{4}{3}+\varepsilon\right)$ approximation for 2VBP **29**
- 4.1 The Analysis of Match&Round . . . 31
  - 4.1.1 A Refined Analysis of the Iterattve Rounding . . . 31
  - 4.1.2 Asymptotic Approximation Ratio . . . 41
  - 4.1.3 Concentration . . . 41
- 4.2 Proof of the Structural Lemma . . . 48
  - 4.2.1 Construction of $\mathcal{S}$ . . . 49
  - 4.2.2 Correctness . . . 51
  - 4.2.3 Refinement for Small Items . . . 59
- 4.3 Existence of ψ-Relaxations . . . 61
- 4.4 Solving the Matching-LP . . . 63

5 Basic Probabilistic Tools **73**

6 Discussion **76**

References **77**

A The Flaw in Bansal, Eliáš and Khan [BEK16] **80**

1 Introduction

Bin Packing is one of the most fundamental problems in combinatorial optimization. An instance of Bin Packing consists of a set I of n items with sizes in $(0, 1]$, for which we seek the smallest number m of unit-size bins into which those items can be packed. The extensive study of Bin Packing since the early 1970's has had a great impact on the design and analysis of approximation algorithms (see, e.g., [FL81, KK82, CCG+13, HR17]).

In this work we study a d-dimensional generalization of Bin Packing, where both the items to be packed as well as bin capacities are given as d-dimensional vectors. Formally, an instance $\mathcal{I}$ of the d-Dimensional Vector Bin Packing (dVBP) problem is a pair (I, v), where I is a set of n items and $v : I \to (0, 1]^d$ is a d-dimensional volume function.[1] A *solution* for the instance (I, v) is a collection of subsets of items $S_1, \ldots, S_m \subseteq I$ such that $v(S_b) = \sum_{i \in S_b} v(i) \leq (1, \ldots, 1)$ for all $b = 1, \ldots, m$ and $\bigcup_{b=1}^{m} S_b = I$.[2] The *size* of the solution is m. Our objective is to find a solution of minimum size.

As a natural generalization of Bin Packing, and due to its wide range of applications, there has been extensive research on dVBP (see, e.g., [GGJY76, FL81, Woe97, KK03, CK04, CHP05, MT06, BCS10, BEK16, ADGH18, WLLH20, Ray21]). Consider, for example, the allocation of computing services (items) to a minimum number of identical servers (bins), where each service requires the use of both CPU and memory. A set of services allocated to a single server may not exceed the available memory and CPU capacity of the server. This yields an instance of 2VBP. For other applications see, e.g., [Spi94, PTUW11, YG12, TS19].

Our goal in this paper is to design efficient polynomial-time approximation algorithms for dVBP. Let $\alpha \geq 1$ be a constant. An algorithm $\mathcal{A}$ is an *asymptotic α-approximation algorithm* for dVBP if for any instance $\mathcal{I}$ of dVBP it returns, in polynomial time, a solution of size at most $\alpha\mathrm{OPT}(\mathcal{I}) + o(\mathrm{OPT}(\mathcal{I}))$, where $\mathrm{OPT}(\mathcal{I})$ is the optimal solution size for $\mathcal{I}$. A weaker notion is that of a *randomized asymptotic α-approximation algorithm*; such an algorithm always returns a solution for $\mathcal{I}$ in polynomial time, but the solution size has to be at most $\alpha\mathrm{OPT}(\mathcal{I}) + o(\mathrm{OPT}(\mathcal{I}))$ with some constant probability. An *asymptotic polynomial-time approximation scheme (APTAS)* is an infinite family $\{\mathcal{A}_\varepsilon\}$ of asymptotic $(1+\varepsilon)$-approximation algorithms, one for each $\varepsilon > 0$. Ray [Ray21] showed that 2VBP does not admit an asymptotic approximation ratio better than $\frac{600}{599}$, assuming $\mathsf{P} \neq \mathsf{NP}$, implying there is no APTAS already for $d = 2$.[3]

In [BCS10] Bansal, Caprara and Sviridenko introduced the *Round&Approx* framework, which yields an asymptotic $(1 + \ln d + \varepsilon)$-approximation for dVBP, for all $d \in \mathbb{N}$ and any $\varepsilon > 0$. Their results are the best-known asymptotic approximation ratio for $d > 3$. For the special cases of $d = 2$ and $d = 3$, the best-known asymptotic approximation ratios, due to Bansal, Eliáš and Khan [BEK16], are $1.5 + \varepsilon$ and $2 + \varepsilon$, respectively, for all $\varepsilon > 0$.

1.1 Our Contribution

Our main contribution is an asymptotic approximation algorithm for dVBP which improves upon the best-known ratio of [BCS10] for all $d > 3$. Specifically, we show the following result.

Theorem 1.1. *For all $d \in \mathbb{N}$ and any $\varepsilon > 0$ there is a randomized $(1 + \ln d - \chi(d) + \varepsilon)$-asymptotic approximation algorithm for dVBP, where* $\chi(d) = \left(\frac{1}{2} \cdot \ln d + \frac{1}{\sqrt{d}} - 1\right) \cdot \left(1 - \sqrt[2d]{\frac{1}{d}}\right)^d > 0.$

Theorem 1.1 is derived via a simple iterative randomized rounding algorithm. In fact, we show that our algorithm outperforms *any* algorithm which follows the framework of Bansal et al. [BCS10].

[1] Instances with $v : I \to [0, 1]^d$ can be easily reduced to equivalent instances with $v : I \to (0, 1]^d$.

[2] We say that $(a_1, \ldots, a_d) \leq (b_1, \ldots, b_d)$ if $a_i \leq b_i$ for $i = 1, \ldots, d$.

[3] Ray's result addresses an oversight in an earlier proof of Woeginger [Woe97].

Reference	$d=2$	$d=3$	$d=4$	arbitrary d
[GGJY76]	2.7	3.7	4.7	$d+0.7$
[FL81]	$2+\varepsilon$	$3+\varepsilon$	$4+\varepsilon$	$d+\varepsilon$
[CK04]				$O(\ln d)$
[KK03]	2 (absolute)			
[BCS10]	≈ 1.69314	≈ 2.09861	≈ 2.38629	$1+\ln d+\varepsilon$
[BEK16]	$\frac{3}{2}+\varepsilon \approx 1.5$ (absolute)	$2+\varepsilon$	$2.5+\varepsilon$	$\frac{d+1}{2}+\varepsilon$
This paper	$\frac{4}{3}+\varepsilon \approx 1.3333$	≈ 2.09801	≈ 2.38617	$1+\ln d-\chi(d)+\varepsilon$

Table 1: Known and new results for dVBP. An entry of value β indicates the paper in this row gives an asymptotic β-approximation for dVBP, where d appears at the head of the column.

For the case $d=2$, we provide a tighter analysis and an additionial *matching* subroutine prior to the iterative randomized rounding phase; together, they enable us to obtain a better bound.

Theorem 1.2. *For any $\varepsilon > 0$, there is a randomized asymptotic $\left(\frac{4}{3}+\varepsilon\right)$-approximation algorithm for 2VBP.*

Table 1 summarizes the previously known, as well as our new results for Vector Bin Packing.

1.2 Related Work

The one-dimensional case (1VBP) is the classic Bin Packing problem. A simple reduction from Partition [Vaz01, Ch. 9] shows there is no α-approximation for Bin Packing with $\alpha < \frac{3}{2}$, assuming $\mathsf{P} \neq \mathsf{NP}$. This motivates the study of asymptotic approximation algorithms for the problem, and in particular, the search for APTASs. The first APTAS for Bin Packing was proposed by Fernandez de la Vega and Lueker [FL81], who introduced the *linear grouping* technique. In their seminal work, Karmarker and Karp [KK82] give an approximation algorithm that uses at most $\mathrm{OPT}(\mathcal{I}) + O(\log^2(\mathrm{OPT}(\mathcal{I})))$ bins. Their work introduced the concept of *Configuration-LP* to which they applied (deterministic) *iterative rounding*. More recently, Hoberg and Rothvoß [HR17] obtained a polynomial-time algorithm that returns a solution of size $\mathrm{OPT}(\mathcal{I}) + O(\log(\mathrm{OPT}(\mathcal{I})))$. Comprehensive surveys of algorithmic results for Bin Packing are given, e.g., by Coffman et al. [CCG+13] and Delorme et al. [DIM16].

To the best of our knowledge, the first asymptotic approximation algorithm for dVBP, due to Garey et al. [GGJY76], achieves the ratio $\left(d+\frac{7}{10}\right)$. This ratio was improved to an asymptotic $(d+\varepsilon)$-approximation by Fernandez de la Vega and Lueker [FL81]. The first algorithm to break the additive of d in the approximation ratio is an asymptotic $(1+O(\ln d))$-algorithm due to Chekuri and Khanna [CK04]. An absolute (i.e., non-asymptotic) 2-approximation ratio for the special case of 2VBP was given by Kellerer and Kotov [KK03].

Bansal, Caprara and Sviridenko [BCS10] introduced a powerful framework, based on randomized rounding, which they call *Round&Approx*. They use it to obtain a randomized asymptotic $(1+\ln d+\varepsilon)$-approximation for dVBP, for every $d \geq 2$ and any $\varepsilon > 0$. The framework combines a *configuration LP* relaxation of the problem with a "subset-oblivious" approximation algorithm.

Informally, a β-subset oblivious algorithm for dVBP is an algorithm which, given a dVBP instance (I, v) and a random subset of items $S \subseteq I$, such that $\Pr(i \in S) \leq \gamma$ for all $i \in I$, returns a solution for (I, v) using approximately $\beta \cdot \gamma \cdot \mathrm{OPT}(I, v)$ bins. A (nearly-optimal) solution for the configuration LP is interpreted as a distribution over the configurations of the instance (i.e., subsets $S \subseteq I$ for which $v_t(S) \leq 1$ for every $t \in \{1, \dots, d\}$). This distribution is used to independently sample a set of configurations; items which do not belong to any of the sampled configurations are packed using the subset-oblivious approximation algorithm. The properties of the subset-oblivious approximation algorithm combined with a concentration bound of McDiarmid [McD89] then yield the claimed approximation guarantee. *Round&Approx* is the framework used to obtain the best approximation algorithms for 2-Dimensional Geometric Bin Packing and for Vector Bin Packing.

Bansal, Eliáš and Khan [BEK16] obtained an asymptotic $\left(\frac{d+1}{2} + \varepsilon\right)$-approximation for dVBP, for all $d \in \mathbb{N}$ and any $\varepsilon > 0$. Their algorithm is based on a rounding scheme which yields a packing with *resource augmentation* in all dimensions except one. The rounding scheme is combined with the generation of an inflated solution of specific structure, which leaves some free volume in all dimensions but one. The free volume is used to balance the resource augmentation. The authors prove the existence of such a solution, while the algorithm uses heavy enumeration to "guess" properties of the solution which suffice to reconstruct it. The authors also attempted to combine this algorithm with the *Round&Approx* framework of Bansal et al. [BCS10] to obtain improved asympatotic approximation. Unfortunately, there is a flaw in the analysis (we give the details in Appendix A).[4] For $d = 2$, Bansal et al. [BEK16] obtained an absolute $(3/2 + \varepsilon)$-approximation using a combinatorial algorithm.

Recently, Sandeep [San22] showed there is no asymptotic $o(\log d)$-approximation for dVBP. For other results relating to dVBP see, e.g., [Joh16] and the excellent survey on multidimensional Bin Packing problems by Christensen et al. [CKPT17].

Iterative rounding and randomized rounding are two powerful techniques used to obtain an integral solution from a fractional solution of an LP relaxation for a problem. Iterative rounding generates an integral solution by iteratively assigning integral values to subsets of variables in the LP, and solving a suitably modified linear program (excluding these variables). In contrast, randomized rounding is done in one shot, by interpreting the variable values as probabilities, and assigning an integral value to each variable via sampling according to these probabilities. An excellent survey on iterative rounding can be found in [LRS11] (see also [Ban14]). For various applications of randomized rounding, see, e.g., [Vaz01, WS11].

One of the earliest and most sophisticated applications of iterative rounding appears in the analysis of Karmarkar and Karp [KK82] in their $\mathrm{OPT}(\mathcal{I}) + O(\log^2(\mathrm{OPT}(\mathcal{I})))$-approximation for classic Bin Packing. Later works applied iterative *randomized* rounding for solving other problems, such as Steiner Tree [BGRS13], makespan minimization on unrelated machines and degree-bounded minimum spanning trees [Ban19], fair scheduling [IM20], and k-Clustering Completion [HS22]. However, we are not aware of earlier use of iterative randomized rounding in solving classic Bin Packing or its variants.

1.3 The Algorithm

Given a dVBP instance (I, v), a *configuration* is a subset $C \subseteq I$ of items such that $v(C) \leq \mathbf{1}$.[5] For each item $i \in I$, let $C(i) \in \{0, 1\}$ indicate whether the item i appears in the configuration C or not. We use $\mathcal{C}$ to denote the set of all configurations. That is, $\mathcal{C} = \{C \subseteq I \mid v(C) \leq (1, \dots, 1)\}$. We use a variant of the standard configuration LP which only consider a subset of items $S \subseteq I$. Given a

[4] We contacted the authors and made them aware of this flaw [BEK21].

[5] We use the notation $\mathbf{1} = (1, \dots, 1)$ and $\mathbf{0} = (0, \dots, 0)$.

Boolean expression $\mathcal{D}$, we define $\mathbb{1}_{\mathcal{D}} \in \{0,1\}$ such that $\mathbb{1}_{\mathcal{D}} = 1$ if $\mathcal{D}$ is true and $\mathbb{1}_{\mathcal{D}} = 0$ otherwise. For every $S \subseteq I$ define

$$\begin{aligned} \mathrm{LP}(S): \quad & \min & & \sum_{C \in \mathcal{C}} \bar{x}_C, \\ & \forall i \in I: & & \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = \mathbb{1}_{i \in S} \\ & \forall C \in \mathcal{C}: & & \bar{x}_C \geq 0 \ . \end{aligned} \tag{1}$$

Each of the variables $\bar{x}_C$ represents a (fractional) selection of the configuration C, where the first constraints ensure that each item $i \in S$ is covered. It is well-known [BCS10] that there is a PTAS for LP(S).

For any vector $\bar{x} \in [0,1]^{\mathcal{C}}$ we associate a distribution over the configurations $\mathcal{C}$. We say that a random configuration $R \in \mathcal{C}$ is *distributed by* $\bar{x}$ (and use the notation $R \sim \bar{x}$) if $\Pr(R = C) = \frac{\bar{x}_C}{z}$ for every $C \in \mathcal{C}$, where $z = \|\bar{x}\| \equiv \sum_{C \in \mathcal{C}} \bar{x}_C$.

Our main algorithm, Iterative Randomized Rounding, is given in Algorithm 1. For arbitrary d, the algorithm is used with $S_0 = I$; the distinction between I and S_0 will be used later in our improved algorithm for 2VBP (see Algorithm 2). We note that Algorithm 1 has a polynomial run time (for fixed δ), and that it returns a solution for the dVBP instance (S_0, v). Line 6 of Algorithm 1 uses a classic First-Fit approach to pack the remaining items (see Section 2 for more details).

Algorithm 1: Iterative Randomized Rounding

Parameters: $\delta \in (0, 0.1)$, $\alpha = -\ln(1-\delta)$ and $k = \lceil \log_{1-\delta}(\delta) \rceil$, where $\delta^{-1} \in \mathbb{N}$.
Input : A d-VBP instance (I, v) and a subset $S_0 \subseteq I$.
Output : A solution for the instance (S_0, v).

1 **for** $j = 1, \ldots, k$ **do**
2   Find a $(1+\delta^2)$-approximate solution $\bar{x}^j$ for LP(S_{j-1}) and let $z_j = \|\bar{x}^j\|$ be its value.
3   Independently sample $\rho_j = \lceil \alpha z_j \rceil$ configurations $C_1^j, \ldots, C_{\rho_j}^j$, where $C_\ell^j \sim \bar{x}^j$ for all $\ell \in \{1, \ldots, \rho_j\}$.
4   Update $S_j \leftarrow S_{j-1} \setminus \left(\bigcup_{\ell=1}^{\rho_j} C_\ell^j \right)$.
5 **end**
6 Pack S_k into configurations $C_1^*, \ldots, C_{\rho^*}^*$ using First-Fit
7 Return $\left(\bigcup_{j=1}^{k} \{C_1^j, \ldots, C_{\rho_j}^j\} \right) \cup \{C_1^*, \ldots, C_{\rho^*}^*\}$.

In the analysis we show that ρ^* is negligible in comparison to OPT(I, v). Thus, the solution generated by Algorithm 1 consists predominately of configurations which are randomly sampled according to solutions for the configuration LP.

Furthermore, the algorithm repeatedly solves the configuration LP, each time using the set S_j consisting of the items not covered in previous iterations. This stands in contrast to algorithms associated with the *Round&Approx* framework (e.g., [BCS10]) which solve the configuration LP once and utilize a subset-oblivious algorithm to generate a significant part of the solution following the random sampling stage.

The above difference is the key for the improved approximation ratio. The analysis of *Round&Approx* uses the fact that if $C_1, \ldots, C_\rho$ are independent random configurations distributed by a (nearly) optimal solution $\bar{x}$ for LP(I), then $\Pr(i \notin \bigcup_{\ell=1}^{\rho} C_\ell) \approx \exp\left(-\frac{\rho}{\mathrm{OPT}(\mathrm{LP}(I))}\right)$. For example, to have the random configurations $C_1, \ldots, C_\rho$ cover each item $i \in I$ (i.e., $i \in C_1 \cup \ldots \cup C_\rho$) with probability $\frac{1}{2}$, the number of sampled configurations has to be $\rho \approx \mathrm{OPT}(\mathrm{LP}(I)) \cdot \ln(2)$. The core idea in our analysis is that if the configurations are sampled iteratively, as in Algorithm 1, then the probability of an item to remain uncovered is $\frac{1}{2}$ after sampling strictly fewer configurations.

Bansal et al. [BCS10] defined the notion of β-subset oblivious algorithms for dVBP; we give a formal definition of the term in Section 3. The main result of Bansal et al. [BCS10], applied to dVBP, is the following.

Theorem 1.3 (*Round&Approx* [BCS10]). *Let $d \in \mathbb{N}$ and $\beta \geq 1$. If there is a polynomial-time β-subset oblivious algorithm for dVBP then there is a randomized asymptotic $(1 + \ln \beta + \varepsilon)$-approximation algorithm for dVBP for every $\varepsilon > 0$.*

Bansal et al. [BCS10] also presented subset-oblivious algorithms for dVBP, as stated in the next lemma.

Lemma 1.4. *For every $\varepsilon > 0$ and $d \in \mathbb{N}$ there is a polynomial-time $(d + \varepsilon)$-subset oblivious algorithm for dVBP.*

In particular, the asymptotic $(1 + \ln d + \varepsilon)$-approximation for dVBP of Bansal et al. [BCS10] is derived as an immediate consequence of Theorem 1.3 and Lemma 1.4. The following theorem states that Algorithm 1 is strictly better than any algorithm that is based on *Round&Approx* (Theorem 1.3).

Theorem 1.5. *Let $\beta \geq 1$ and $d \in \mathbb{N}$. If there is a β-subset oblivious algorithm for dVBP then for every $\varepsilon > 0$ there exists $\delta > 0$ such that Algorithm 1 configured with δ is a randomized asymptotic $(1 + \ln \beta - \chi(\beta, d) + \varepsilon)$-approximation algorithm for dVBP, where $\chi(\beta, d) = \left(\frac{1}{2} \cdot \ln \beta + \frac{1}{\sqrt{\beta}} - 1\right) \cdot \left(1 - \sqrt[2d]{\frac{1}{\beta}}\right)^d$.*

Theorem 1.1 follows immediately from Theorem 1.5 and Lemma 1.4. Since $\chi(\beta, d) > 0$ for all $\beta > 1$ and $d > 1$, Theorem 1.5 implies that Algorithm 1 is strictly better than *Round&Approx*; that is, it achieves an asymptotic approximation ratio smaller than that obtained by any *Round&Approx*-based algorithm. Furthermore, while the result of Theorem 1.3 refers to an algorithm which uses as a subroutine a β-subset oblivious algorithm, the result of Theorem 1.5 uses the β-subset oblivious algorithm only as part of its proof. Thus, Theorem 1.5 does not require the subset-oblivious algorithm to run in polynomial time. Finally, we note that the value of $\chi(\beta, d)$ in Theorem 1.5 is likely to be sub-optimal, and can probably be replaced by a larger value. Our main objective is to show Algorithm 1 yields a better asymptotic approximation ratio in comparison to *Round&Approx* in a simple manner, possibly sacrificing the value of $\chi(\beta, d)$.

The proof of Theorem 1.5 utilizes an iteration-dependent bound on $\mathrm{OPT}(S_j, v)$. Trivially, $\mathrm{OPT}(S_j, v) \leq \mathrm{OPT}(I, v)$. The subset-oblivious algorithm is used to show that $\mathrm{OPT}(S_j, v) \lesssim \beta(1-\delta)^j \mathrm{OPT}(I, v)$ with high probability. Together, these two bounds can be used to show that the asymptotic approximation ratio of Algorithm 1 is approximately $(1 + \ln \beta)$, matching the statement of Theorem 1.3. To show a strictly better approximation ratio, we consider a nearly optimal solution $A_1, \ldots, A_m$ of the instance, and use a simple rounding scheme to show that if $T_j \subseteq \{1, \ldots, m\}$ is a set of configurations such that $v_t(A_b \cap S_j) \leq 1 - \delta$ for all $t = 1, \ldots, d$, then the items in $\left(\bigcup_{b \in T_j} A_b\right) \cap S_r$ can be packed in strictly less than $|T_j|$ configurations for $r > j$. This, together with a lower bound on $|T_j|$ for a specific iteration j, leads to a third upper bound on $\mathrm{OPT}(S_j, v)$, which is used to obtain the improved asymptotic approximation ratio. The configurations in T_j can be considered as "easy", and the lower bound on $|T_j|$ can be interpreted as a guarantee that some configurations must "become easy" as the iterative rounding process progresses.

In Section 3 we further show that the dependence of δ on ε derived from Theorem 1.5 is polynomial. A simple consequence of this property is that, by appropriately setting δ, Algorithm 1 is a randomized asymptotic $(1+\varepsilon)$-approximation for Bin Packing whose run time is polynomial in the input size and in $\frac{1}{\varepsilon}$. Thus, we have

Lemma 1.6. *Algorithm 1 is a randomized asymptotic fully polynomial-time approximation scheme (AFPTAS)*[6] *for* Bin Packing.

1.4 Improved Algorithm for 2VBP

For the special case where $d = 2$, we strengthen our analysis to obtain a better approximation ratio. To simplify our analysis, we may assume our instances adhere to a specific structure. Given $\delta > 0$, we say that an item $i \in I$ is δ-*huge* if $v_1(i) \geq 1-\delta$ and $v_2(i) \geq 1-\delta$. The δ-huge free 2VBP (δ-2VBP) is the special case of 2VBP in which there are no δ-huge items. In solving a general 2VBP instance, we may restrict our attention to the corresponding δ-huge free instance, as formalized in the next result.

Lemma 1.7. *For any $\alpha \geq 1$ and $\delta \in (0, 0.1)$, if there is a randomized asymptotic α-approximation for δ-2VBP then there is a randomized asymptotic $(\alpha + 4\delta)$-approximation for 2VBP.*

The lemma follows by noting that each huge item can be packed in a separate bin. This incurs only a small increase in the packing size (we omit the details).

The analysis of Algorithm 1 (as part of our approximation algorithm for 2VBP) relies on an iteration-dependent bound on $\mathrm{OPT}(S_j, v)$ which holds with high probability. We use a classification of items and configurations into categories. As in Algorithm 1, let $\delta \in (0,1)$ be such that $\delta^{-1} \in \mathbb{N}$. We say that an item $i \in I$ is δ-*large* if $v_1(i) > \delta$ or $v_2(i) > \delta$, and use $L \subseteq I$ to denote the set of δ-large items (δ is commonly known by context). It can be easily shown that $|C \cap L| \leq 2 \cdot \delta^{-1}$ for all $C \in \mathcal{C}$. For $h = 2, \ldots, 2 \cdot \delta^{-1}$, we define

$$\mathcal{C}_h = \{C \in \mathcal{C} \mid v(C \cap L) > (1-\delta, 1-\delta) \text{ and } |C \cap L| = h\}. \tag{2}$$

Let $\mathcal{C}_0 = \mathcal{C} \setminus \left(\bigcup_{h=2}^{2\cdot\delta^{-1}} \mathcal{C}_h \right)$ be the set of all remaining configurations. As we assume that (I, v) is an instance of δ-2VBP (i.e., no δ-huge items), it follows that for every $C \in \mathcal{C}_0$ either $v_1(C \cap L) \leq 1-\delta$ or $v_2(C \cap L) \leq 1-\delta$.

For vectors $\bar{x}, \bar{z} \in [0,1]^{\mathcal{C}}$, $\bar{x} \cdot \bar{z} = \sum_{C \in \mathcal{C}} \bar{x}_C \cdot \bar{z}_C$ is the dot product of $\bar{x}$ and $\bar{z}$. By applying a tigher analysis (in comparison to Theorem 1.5), it can be shown that if $\bar{x}^* \in [0,1]^{\mathcal{C}}$ is a solution for $\mathrm{LP}(S_0)$ then,with high probability, the solution returned by Algorithm 1 is of size at most

$$\bar{x}^* \cdot \mathbb{1}_{\mathcal{C}_0} + \sum_{h=2}^{2\delta^{-1}} \frac{h+1}{h} \cdot \bar{x}^* \cdot \mathbb{1}_{\mathcal{C}_h} \leq \left(\frac{3}{2} + \delta \cdot O(1) \right) \cdot \|\bar{x}^*\| + O(1) \ . \tag{3}$$

This implies that, given the input $S_0 = I$, and by taking $\bar{x}^*$ which corresponds to an optimal solution, Algorithm 1 yields an asymptotic approximation ratio arbitrarily close to $\frac{3}{2}$. While we do not include a proof of (3), the proof can be derived by modifying the proof of Lemma 4.11 and using Lemma 4.16.

Our analysis relies on structural properties of 2VBP instances (inspired by properties presented by Bansal et al. [BEK16]) by which configurations in $\mathcal{C}_0$ are "easy" (when selected by $\bar{x}^*$) and configuration in $\mathcal{C} \setminus \mathcal{C}_0$ are "difficult". Intuitively, from the viewpoint of Algorithm 1, a configuration

[6] A randomized AFPTAS for a problem $\mathcal{P}$ is an infinite family $\{\mathcal{A}_\varepsilon\}$ of randomized asymptotic $(1+\varepsilon)$-approximation algorithms for $\mathcal{P}$, one for each $\varepsilon > 0$, whose run times are polynomial in the input size and in $\frac{1}{\varepsilon}$.

$C \in \mathcal{C} \setminus \mathcal{C}_0$ *becomes* easy at iteration j if $C \cap L \not\subseteq S_j$, as in this case $C \cap S_j \in \mathcal{C}_0$. Our analysis exploits this intuition via the notion of *touched* and *untouched* configurations (see the formal definition in Section 4.1.1).

The bound (3) on the solution quality suggests that the most "difficult" configurations in $\bar{x}^*$ are those in $\mathcal{C}_2$; indeed, if we have an optimal solution containing no configuration in $\mathcal{C}_2$ then we can obtain an approximation ratio of $\frac{4}{3}$. Furthermore, if an optimal (integral) solution contains only configurations in $\mathcal{C}_2$ then a nearly optimal solution can be easily constructed using *matching*. As a solution may contain both configurations in $\mathcal{C}_2$ and in $\mathcal{C} \setminus \mathcal{C}_0$, we use a sophisticated combination of a matching polytope and a configuration LP, along with the dependent sampling technique of Chekuri et al. [CVZ11]. In the execution of our algorithm Match&Round, the solution for the resulting LP is (conceptually) partitioned into two parts: one which contains the configurations in $\mathcal{C}_2$ and handled using matching techniques, and another which contains the remaining configurations that is handled by Algorithm 1.

We define the *δ-matching graph* $G = (L, E)$ of (I, v) as the graph whose vertex set L consists of the δ-large items of (I, v), and whose edge set is $E = \{\{i_1, i_2\} \subseteq L \mid \{i_1, i_2\} \in \mathcal{C}_2\}$. We use $P_{\mathcal{M}}(G)$ to denote the matching polytope of G. We refer the reader to Schrijver's book [Sch03] for a formal definition of the matching polytope. Given $\bar{x} \in [0,1]^{\mathcal{C}}$, we define the projection of $\bar{x}$ on E as the vector $\bar{p} \in \mathbb{R}^E_{\geq 0}$ where $\bar{p}_e = \sum_{C \in \mathcal{C} \text{ s.t. } e \subseteq C} \bar{x}_C$. Let $\mathcal{E}(\bar{x}) = \bar{p}$. We note that for any $C \in \mathcal{C}$, there is at most a single edge $e \in E$ such that $e \subseteq C$.

The *Matching Configuration LP* of the δ-2VBP instance (I, v) is the following optimization problem:

$$\begin{aligned} \text{MLP}: \quad & \min & & \sum_{C \in \mathcal{C}} \bar{x}_C, \\ & \forall i \in I: & & \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = 1, \\ & & & \mathcal{E}(\bar{x}) \in P_{\mathcal{M}}(G), \\ & \forall C \in \mathcal{C}: & & \bar{x}_C \geq 0 \ . \end{aligned} \tag{4}$$

Thus, MLP takes as input a δ-2VBP instance (I, v), and a *solution* for (I, v) is a vector $\bar{x} \in \mathbb{R}^{\mathcal{C}}_{\geq 0}$ which satisfies the constraints in (4). The objective is to find a solution $\bar{x}$ such that $\|\bar{x}\| = \sum_{C \in \mathcal{C}} \bar{x}_C$ is minimized.

Note that the Matching Configuration LP is a restriction of LP(I) in which we also require that $\mathcal{E}(\bar{x})$ is in the matching polytope $P_{\mathcal{M}}(G)$. Observe that if $S_1, \ldots, S_m$ is a solution for (I, v) in which the sets $S_1, \ldots, S_m$ are pairwise disjoint, then the vector $\bar{x} \in \{0,1\}^{\mathcal{C}}$ with $\bar{x}_{S_b} = 1$ for $b \in \{1, \ldots, m\}$ and $\bar{x}_C = 0$ for any other $C \in \mathcal{C}$, is a feasible solution for MLP. This holds since the set $\{e \in E \mid \exists b \in \{1, \ldots, m\}: \ e \subseteq S_b\}$ forms a matching in the graph G.

Similar to the configuration LP, MLP can be approximated as well:

Lemma 1.8. *For any $\delta \in (0, 0.1)$, there is a PTAS for the* MLP *problem.*

We note that writing $P_{\mathcal{M}}(G)$ as a linear program requires a super-polynomial number of constraints [Rot17]. It follows that both MLP and its dual have super-polynomial number of variables and a super-polynomial number of constraints. Thus, the standard method for solving configuration LPs using an approximate separation oracle for the dual program fails (the method can be traced back to Karmarker and Karp [KK82]), and more sophisticated tools are required to obtain a PTAS. We give the proof of Lemma 1.8 in Section 4.4.

Given $\bar{x}$ such that $\bar{\beta} = \mathcal{E}(\bar{x}) \in P_{\mathcal{M}}(G)$ and a parameter $\gamma > 0$, we use a randomized algorithm of Chekuri, Vondrák and Zenklusen [CVZ11] called SampleMatching. This algorithm, for input $(\bar{\beta}, \gamma)$

in polynomial time generates a random matching $\mathcal{M}$ for which $\Pr(e \in \mathcal{M}) = (1-\gamma)\bar{\beta}_e$. Importantly, the algorithm also gives dimension-free Chernoff-like concentration bounds for $\mathcal{M}$ (see Lemma 4.20 for details).

We refer to our algorithm for 2VBP as Match&Round; its pseudocode is given in Algorithm 2.[7] We note that Match&Round is a polynomial-time algorithm which returns a solution for the instance (I, v).

Algorithm 2: Match&Round

Parameters: $0 < \delta < 0.1$, where $\delta^{-1} \in \mathbb{N}$.
Input : A δ-2VBP instance (I, v).
Output : A solution for the instance (I, v).

1 Find a $(1 + \delta^2)$-approximate solution $\bar{x}^0$ for MLP.
2 $\mathcal{M} \leftarrow \mathsf{SampleMatching}\left(\mathcal{E}(\bar{x}^0), \delta^4\right)$, and set $S_0 \leftarrow I \setminus \left(\bigcup_{e \in \mathcal{M}} e\right)$.
3 Run Algorithm 1 on the instance (I, v) with S_0 and the parameter δ. Denote the returned solution by $D_1, \ldots, D_m$.
4 Return $\mathcal{M} \cup \{D_1, \ldots, D_m\}$.

Our main result for 2VBP follows from the next lemma.

Lemma 1.9. *For any $\delta \in (0, 0.1)$, Algorithm 2 is a randomized asymptotic $\left(\frac{4}{3} + O(\delta)\right)$-approximation for δ-2VBP.*

Using Lemma 1.9 and Lemma 1.7, we obtain the statement of Theorem 1.2. We use the standard notation of $\wedge$ for the element-wise minimum of two vectors.[8] The analysis of Algorithm 2 is based on a partition of the solution $\bar{x}^0$ obtained in Line 2 into its two "matching" and "fractional" components: $\bar{x}^0 \wedge \mathbb{1}_{\mathcal{C}_2}$ and $\bar{x}^0 \wedge \mathbb{1}_{\mathcal{C} \setminus \mathcal{C}_2}$. We show that, with high probability, $|\mathcal{M}| \lesssim \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C}_2}$. Furthermore, we exploit the fact that $\bar{x}^0 \wedge \mathbb{1}_{\mathcal{C} \setminus \mathcal{C}_2}$ does not select configurations in $\mathcal{C}_2$ to show that the number of configurations returned by Algorithm 1 (when invoked in Step 3 of Algorithm 2) is bounded by $\approx \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C} \setminus \mathcal{C}_2} + \frac{1}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C}_2}$.

1.5 Technical Contribution

Our main technical contribution is the introduction of iterative *randomized* rounding in the context of Bin Packing. The ingenious randomized rounding techniques known for Bin Packing problems (e.g., [BCS10]) rely on solving *once* a Configuration-LP and sampling a set of configurations according to the distribution induced by the Configuration-LP solution. In contrast, our iterative randomized rounding approach is based on solving a (modified) Configuration-LP iteratively and sampling in each iteration a set of configurations using the distribution induced by the current LP solution. While the resulting algorithms are simple and yield improved ratios, we are not aware of the use of iterative randomized rounding in previous studies of Bin Packing problems.

Intuitively, we expect iterative randomized rounding to outperform non-iterative randomized rounding in the context of Bin Packing. Indeed, the former is less likely to select many configurations containing the same item, presumably leading to a more efficient solution. Moreover, once a significant fraction (say, 10%) of the items in I is "covered" (at random), we expect the Configuration-LP solution value to decrease. This can be used to obtain a better approximation ratio if we solve the *modified* Configuration-LP, and use the corresponding distribution to sample

[7] The idea to use matching algorithms is inspired by the work of Bansal et al. [BEK16]. However, matching plays different roles in the two algorithms. In particular, MLP is introduced in this paper.

[8] That is, for $\bar{r}^1 = (\bar{r}^1_1, \ldots, \bar{r}^1_k)$ and $\bar{r}^2 = (\bar{r}^2_1, \ldots, \bar{r}^2_k)$, $(\bar{r}^1 \wedge \bar{r}^2)_i = \min\{\bar{r}^1_i, \bar{r}^2_i\}$ for $i = 1, \ldots, k$.

configurations. However, formalizing the above intuition into a rigor proof is non-trivial. In the proof of Theorem 1.5 we provide a formal expression to the above intuition and prove that iterative randomized rounding is superior to any algorithm which follows the *Round&Approx* framework.

Our analysis for the case of arbitrary $d > 2$ is fairly simple, leaving much room for improvement. For the special case of $d = 2$ we use a tighter analysis. While many of the ingredients in this tighter analysis can be applied also to dVBP instances where $d > 2$, and possibly to other Bin Packing variants, some of the concepts exploit special properties of 2VBP instances. This includes a strong structural property (Lemma 4.2) on which we elaborate in Section 4, and the Matching-Configuration-LP. The strong structural property can potentially be extended to the d-dimensional case; however, such extension requires overcoming some technical challenges. We elaborate on these challenges in Section 6.

1.6 Organization

In Section 2 we give some definitions and notation. Section 3 presents the analysis of Algorithm 1 as well as the proof of Theorem 1.5 and Lemma 1.6. Section 4 gives the results for 2VBP including the PTAS for the Matching-Configuration-LP (Lemma 1.8). In Section 5 we show basic properties which are used both in Sections 3 and 4. We conclude with a discussion in Section 6.

2 Preliminaries

In this section we give some basic definitions and properties that will be used in the proofs of Theorem 1.5 and Theorem 1.2. Throughout the paper, for $x \in \mathbb{R}$, $\exp(x) = e^x$, where $e = 2.718...$ is the base of the natural logarithm.

2.1 Probability Space

Our analysis refers to an execution of either Algorithm 1 or Algorithm 2 on a dVBP instance (I, v). For an execution of Algorithm 1, we have $S_0 = I$. We use $(\Omega, \mathcal{F}, \Pr)$ to denote the probability space generated by the algorithm. Observe that as $\delta < 0.1$,

$$\rho_j \leq \lceil \alpha z_j \rceil \leq \left\lceil (-\ln(1-\delta))\,(1+\delta^2)\mathrm{OPT} \right\rceil \leq \mathrm{OPT}, \qquad \text{for } j = 1, \ldots, k \ .$$

Assume, without loss of generality, that Algorithm 1 samples in each iteration OPT configurations $C_1^j, \ldots, C_{\mathrm{OPT}}^j$ independently according to $\bar{x}^j$, and ignores configurations $C_{\rho_j+1}^j, \ldots, C_{\mathrm{OPT}}^j$. Furthermore, we may assume that Ω is finite. Define the random variables $P_0 = S_0$ and $P_j = (C_1^j, \ldots, C_{\mathrm{OPT}}^j)$ for $j = 1, \ldots, k$. Let $\mathcal{F}_j = \sigma(P_0, P_1, \ldots, P_j)$ be the σ-algebra of the random variables $P_0, P_1, \ldots, P_j$. We also define $\mathcal{F}_{-1} = \{\emptyset, \Omega\}$. It follows that $\mathcal{F}_{-1} \subseteq \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_k$.

We use conditional expectations and probabilities given the σ-algebra $\mathcal{F}_j$. We refer the reader to standard textbooks on probability (e.g., by Chow and Teicher [CT97]) for the formal definitions. Intuitively, $\mathbb{E}[X|\mathcal{F}_j]$ is the expectation of X given the sample outcomes up to iteration j, and as such depends on the outcomes of the first j iterations.

The parameter α is set such that the probability of $i \in S_j$ decreases exponentially with j, as stated in the next lemma.

Lemma 2.1. *For $j = 1, \ldots, k$ and $i \in I$ it holds that*

$$\Pr(i \in S_j \mid \mathcal{F}_{j-1}) = \mathbb{1}_{i \in S_{j-1}} \cdot \left(1 - \frac{1}{z_j}\right)^{\rho_j} \leq (1-\delta) \cdot \mathbb{1}_{i \in S_{j-1}} \ .$$

Proof. We can write

$$\Pr\left(\mathbb{1}_{i\in S_j}\middle|\mathcal{F}_{j-1}\right) = \mathbb{1}_{i\in S_{j-1}}\cdot \Pr\left(\forall \ell \in \rho_j:\ i\notin C^j_\ell \,\middle|\, \mathcal{F}_{j-1}\right) = \mathbb{1}_{i\in S_{j-1}}\cdot \prod_{\ell=1}^{\rho_j}\Pr\left(i\notin C^j_\ell \,\middle|\, \mathcal{F}_{j-1}\right)$$
$$= \mathbb{1}_{i\in S_{j-1}}\left(1-\frac{\mathbb{1}_{i\in S_{j-1}}}{z_j}\right)^{\rho_j} = \mathbb{1}_{i\in S_{j-1}}\left(1-\frac{1}{z_j}\right)^{\rho_j} \leq \mathbb{1}_{i\in S_{j-1}}\cdot \exp\left(-\alpha\right) = \mathbb{1}_{i\in S_{j-1}}\cdot (1-\delta)\ . \tag{5}$$

The first equality holds by the definition of S_j, and the second holds since $C^j_1,\dots,C^j_{\rho_j}$ are conditionally independent given $\mathcal{F}_{j-1}$ (note that ρ_j is $\mathcal{F}_{j-1}$-measurable). The third equality holds since $\bar{x}^j$ is a solution for $\mathrm{LP}(\mathbb{1}_{S_{j-1}})$ and $C^j_\ell \sim \bar{x}^j$. The inequality in (5) uses $\rho_j \geq \alpha z_j$ and $\left(1-\frac{1}{x}\right)^x \leq \exp(-1)$ for $x \geq 1$. ☐

2.2 McDiarmid's Concentration Bound

Our analysis heavily relies on concentration bounds. Let A be an arbitrary set, $m \in \mathbb{N}_+$ and $f: A^m \to \mathbb{R}$. For any $\eta \geq 0$, we say that f is of *η-bounded difference* if for any $\bar{x}, \bar{x}' \in A^m$ and $r \in \{1,\dots,m\}$ such that $\bar{x}_\ell = \bar{x}'_\ell$ for all $\ell \in \{1,\dots,m\}\setminus\{r\}$ (i.e., $\bar{x}$ and $\bar{x}'$ differ only in the r-th entry) it holds that $|f(\bar{x}) - f(\bar{x}')| \leq \eta$. The next result is due to McDiarmid [McD89].

Lemma 2.2 (McDiarmid)**.** *Given a finite arbitrary set A, $m \in \mathbb{N}_+$ and $\eta > 0$, let $f: A^m \to \mathbb{R}$ be a function of η-bounded difference. Also, let $X_1,\dots,X_m \in A$ be independent random variables. Then for any $t \geq 0$,*

$$\Pr\left(f(X_1,\dots,X_m) - \mathbb{E}\left[f(X_1,\dots,X_m)\right] > t\right) \leq \exp\left(-\frac{2\cdot t^2}{m\cdot \eta^2}\right)\ .$$

To motivate our next lemma, consider the following example arising in our setting. Let $x_0, x_1, \dots, x_k$ be random variables defined by $x_j = \sum_{t=1}^d v_t(S_j)$. That is, x_j is the total volume of S_j in all dimensions. Given S_{j-1} and ρ_j we we can express x_j as a function of $C^j_1,\dots,C^j_{\mathrm{OPT}}$. For any $S\subseteq I$ and $\rho \in [\mathrm{OPT}]$ define $f_{S,\rho}: \mathcal{C}^{\mathrm{OPT}} \to \mathbb{R}$ by

$$f_{S,\rho}(C_1,\dots,C_{\mathrm{OPT}}) = \sum_{t=1}^d v_t\left(S\setminus\left(\bigcup_{\ell=1}^{\rho} C_j\right)\right)\ .$$

Then it can be verified that $x_j = g(C^j_1,\dots,C^j_{\mathrm{OPT}})$ where $g = f_{S_{j-1},\rho_j}$. However, we cannot use Lemma 2.2 to show that $x_j \approx \mathbb{E}[x_j]$ with high probability, since the random variables $C^j_1,\dots C^j_{\mathrm{OPT}}$ are not independent, and the function g is random.

Nontheless, we note that at the end of iteration $j-1$ (Step 1 of Algorithm 1) the values of S_{j-1} and ρ_j are known (while ρ_j was not computed yet, its value does not depend on future random samples); thus, the function $g = f_{S_{j-1},\rho_j}$ is known at iteration j of the algorithm. Furthermore, the random variables $C^j_1,\dots,C^j_{\mathrm{OPT}}$ are independent (by definition) assuming we have the random samples of the first $(j-1)$ iterations. Therefore, we expect Lemma 2.2 to hold in this setting. More formally, since $C^j_1,\dots,C^j_{\mathrm{OPT}}$ are conditionally independent[9] given $\mathcal{F}_{j-1}$, and $g = f_{S_{j-1},\rho_j}$ is a random function that is $\mathcal{F}_{j-1}$-measurable, we expect that $g(C^j_1,\dots,C^j_{\mathrm{OPT}}) \approx \mathbb{E}[g(C^j_1,\dots,C^j_{\mathrm{OPT}})|\mathcal{F}_{j-1}]$. This is formalized in the next lemma.

[9] See, e.g., the book by Chow and Teicher [CT97] for a formal definition of conditional independence.

Lemma 2.3 (Generalized McDiarmid)**.** *Given a finite arbitrary set A, $m \in \mathbb{N}_+$ and $\eta > 0$, let D be a finite family of η-bounded difference functions from A^m to $\mathbb{R}$. Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space for which Ω is finite, $\mathcal{G} \subseteq \mathcal{F}$ a σ-algebra, and $g \in D$ a $\mathcal{G}$-measurable random function (i.e., $g : \Omega \to D$ with $\{\omega \in \Omega |\ g(\omega) \in U\} \in \mathcal{G}$ for every $U \subseteq D$). Then, for a sequence of random variables $X_1, \ldots, X_m \in A$ which are conditionally independent given $\mathcal{G}$, and any $t \geq 0$,*

$$\Pr\left(g(X_1, \ldots, X_m) - \mathbb{E}\left[g(X_1, \ldots, X_m) | \mathcal{G}\right] > t\right) \leq \exp\left(-\frac{2 \cdot t^2}{m \cdot \eta^2}\right) .$$

Lemma 2.3 can be derived from Lemma 2.2 using standard arguments from probability theory (we omit the details).

We use Lemma 2.3 in the proofs of Theorems 1.5 and 1.2. For a set of items $S \subseteq I$, we denote by $\mathbb{1}_S \in \{0,1\}^I$ an indicator vector in which entries corresponding to $i \in S$ are equal to '1', and all other entries are equal to '0'.[10] The next lemma is used in the proofs of both theorems, and deals with random variables of the form $\mathbb{1}_{S_j} \cdot \bar{u}$ where $\bar{u} \in \mathbb{R}^I_{\geq 0}$. Given $\bar{u} \in \mathbb{R}^I$ define the *tolerance* of $\bar{u}$ by $\mathsf{tol}(\bar{u}) = \max_{C \in \mathcal{C}} \left(\sum_{i \in C} \bar{u}_i\right)$. Intuitively, the vector $\bar{u}$ associates with each item $i \in I$ some weight $\bar{u}_i$; then $\mathsf{tol}(\bar{u})$ is the largest total weight of a configuration C with respect to $\bar{u}$.

Lemma 2.4. *Let $j \in \{0, 1, \ldots, k-1\}$ and $t > 0$. Also, let $\bar{u} \in \mathbb{R}^I_{\geq 0}$ be an $\mathcal{F}_j$-measurable random vector. Then,*

$$\Pr\left(\exists r \in \{j, \ldots, k\} :\ \bar{u} \cdot \mathbb{1}_{S_r} - (1-\delta)^{r-j} \cdot \bar{u} \cdot \mathbb{1}_{S_j} > t \cdot \mathsf{tol}(\bar{u})\right) \leq \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^4 \cdot t^2}{\mathrm{OPT}}\right) .$$

The proof of Lemma 2.4, given in Section 5, follows from Lemma 2.3 and Lemma 2.1.

2.3 First-Fit

In several places we use the following First-Fit strategy, which takes as input a dVBP instance (I, v) and a subset of items $S \subseteq I$. Throughout its execution, First-Fit maintains a set $A_1, \ldots, A_m \subseteq S$ of configurations, and iterates over the items in S. For each item $i \in S$, First-Fit examines the configurations sequentially, until it finds a configuration A_i to which i can be added without violating the volume constraints. If no such configuration exists, First-Fit adds a new configuration $A_{m+1} = \{i\}$. The next lemma follows from a simple analysis of First-Fit for Bin Packing (see, e.g., Vazirani [Vaz01, Ch. 9]), by taking for each item $i \in I$ in the dVBP instance $\hat{v}(i) = \max\{v_1(i), \ldots, v_d(i)\}$, and considering the problem in single dimension.

Lemma 2.5. *Given a dVBP instance (I, v) and a subset of items $S \subseteq I$, First-Fit returns a packing of S in at most $2 \cdot \left(\sum_{t=1}^{d} v_t(S)\right) + 1$ bins.*

Recall that ρ^* is the number of configurations used by the First-Fit strategy in Step 6 of Algorithm 1. By Lemma 2.1, it follows that $\mathbb{E}\left[\sum_{t=1}^{d} v_t(S_k)\right] \leq (1-\delta)^k \left(\sum_{t=1}^{d} v_t(I)\right) \leq d \cdot \delta \mathrm{OPT}$, and by Lemma 2.5 we have $\mathbb{E}[\rho^*] \leq 2 \cdot d \cdot \delta \mathrm{OPT} + 1$. The next lemma uses Lemma 2.4 to show that, with high probability, ρ^* does not significantly deviate from its expectation.

Lemma 2.6. *With probability at least $1 - \delta^{-2} \cdot \exp\left(-\delta^7 \cdot \mathrm{OPT}\right)$, it holds that $\rho^* \leq 8 \cdot d \cdot \delta \cdot \mathrm{OPT} + 1$.*

The proof of Lemma 2.6 is given in Section 5. Lemma 2.6 implies that the number of configurations added by the First-Fit strategy in Line 6 of Algorithm 1 is negligible.

[10] Similarly, for a set of configurations $\mathcal{C}' \in \mathcal{C}$, we use the indicator vector $\mathbb{1}_{\mathcal{C}'} \in \{0,1\}^{\mathcal{C}}$ in which entries corresponding to $C \in \mathcal{C}'$ are equal to '1'.

3 Improved Asymptotic Approximation for dVBP

In this section we prove Theorem 1.5. That is, we show that Algorithm 1 outperforms *any* algorithm which falls into the *Round&Approx* framework of Bansal et al. [BCS10]. We also derive Lemma 1.6 as a simple consequence of the analysis of Algorithm 1.

As Theorem 1.5 refers to subset-oblivious algorithms, we first have to formally define this class of algorithms. The following is a slight simplification of the definition of Bansal et al. [BCS10, Definition 1].

Definition 3.1. *For every $d \in \mathbb{N}$ and $\beta \geq 1$, an algorithm* appr *is β-*subset oblivious *for dVBP if for every $\varepsilon > 0$ there are $K \in \mathbb{N}$ and $\psi > 0$ such that, for every dVBP instance (I, v), there is a set of K vectors $\mathcal{S} \subseteq \mathbb{R}^I_{\geq 0}$ which satisfies the following properties:*

1. *For any $\bar{u} \in \mathcal{S}$, it holds that $\mathsf{tol}(\bar{u}) \leq \psi$.*
2. $\mathrm{OPT}(I, v) \geq \max_{\bar{u} \in \mathcal{S}} \|\bar{u}\|$.
3. *For any $Q \subseteq I$, given the dVBP instance (Q, v),* appr *returns a solution satisfying*

$$\mathsf{appr}(I, v, Q) \leq \beta \cdot \max_{\bar{u} \in \mathcal{S}} \mathbb{1}_Q \cdot \bar{u} + \varepsilon \cdot \mathrm{OPT}(I, v) + K,$$

where $\mathsf{appr}(I, v, Q)$ is the number of bins used by the solution.

*We refer to K and ψ as the ε-*parameters of appr*, and to $\mathcal{S}$ as the ε-*weight vectors of appr and *(I, v).*

Instead of Theorem 1.5 we prove a more specific result, which indicates also the dependencies between ε and δ.

Theorem 3.2. *Let $\beta \geq 1$ and $d \in \mathbb{N}$. If there is a β-subset oblivious algorithm for dVBP then for every $\delta \in (0, 0.1)$ such that $\delta^{-1} \in \mathbb{N}$ and $\delta < \min\left\{\frac{1}{28d^2}, \frac{1}{\beta}\right\}$ it holds that Algorithm 1 configured with δ is a randomized asymptotic $(1 + \ln\beta - \chi(\beta, d) + 200 \cdot d^2 \cdot \delta \cdot \beta)$-approximation algorithm for dVBP, where $\chi(\beta, d) = \left(\frac{1}{2} \cdot \ln\beta + \frac{1}{\sqrt{\beta}} - 1\right) \cdot \left(1 - \sqrt[2d]{\frac{1}{\beta}}\right)^d$.*

We give the proof of Theorem 3.2 in Section 3.1. We first use Theorem 3.2 to derive Lemma 1.6.

Proof of Lemma 1.6. Let $\varepsilon \in (0, 0.1)$ and $\delta = \frac{1}{\lceil 400 \cdot \varepsilon^{-1} \rceil}$. Consider the execution of Algorithm 1 with a Bin Packing (1VBP) instance and the above parameter δ. By Lemma 1.4 there is a $(1 + \delta)$-subset oblivious algorithm for Bin Packing; thus, by Theorem 3.2, Algorithm 1 is a randomized asymptotic ζ-approximation for Bin Packing, where

$$\zeta \;=\; (1 + \ln(1 + \delta) + 200 \cdot \delta \cdot (1 + \delta) - \chi(1 + \delta, 1)) \;\leq\; (1 + \delta + 200 \cdot \delta \cdot (1 + \delta)) \;\leq\; (1 + \varepsilon)\ .$$

The first inequality holds as $\chi(1 + \delta, 1) \geq 0$ and $\ln(1 + \delta \leq \delta)$. The last inequality follows from $400 \cdot \delta \leq \varepsilon$ by the definition of δ.

It is well-known that (1) admits an FPTAS for Bin Packing instances. Indeed, in this case the separation oracle for the dual of (1) needs to solve an instance of the (standard) Knapsack problem, for which there is an FPTAS (see, e.g., Vazirani's textbook [Vaz01]). Hence, the run time of each iteration in Line 1 of Algorithm 1 (given a Bin Packing instance) is polynomial in the instance size and δ^{-2}. As the total number of iterations is $k \leq \delta^{-2}$, it follows that the total run time is polynomial in the input size and in $\frac{1}{\delta}$. Since we defined δ to be polynomial in ε, it follows that the run time is polynomial in the input size and $1/\varepsilon$. Thus, Algorithm 1 is an AFPTAS for Bin Packing. ∎

3.1 Proof of Theorem 3.2

Let (I, v) be a dVBP instance, and let appr be a β-subset oblivious algorithm for dVBP. Also, let $\delta \in (0, 0.1)$ such that $\delta \le \frac{1}{28 \cdot d^2}$, $\delta < \frac{1}{\beta}$, and $\delta^{-1} \in \mathbb{N}$. We denote by $\mathrm{OPT} = \mathrm{OPT}(I, v)$ the value of an optimal solution for the instance. Consider an execution of Algorithm 1 with the instance (I, v), $S_0 = I$ and the parameter δ. We use notations such as ρ_j, S_j and C_ℓ^j when referring to the corresponding variables in the execution of Algorithm 1. We also use the probability space $(\Omega, \Pr, \mathcal{F})$ and the filtration $\mathcal{F}_{-1}, \mathcal{F}_0, \ldots, \mathcal{F}_k$ as defined in Section 2.

The size of the solution returned by Algorithm 1 is $\sum_{j=1}^{k} \rho_j + \rho^*$. By Lemma 2.6, the value of ρ^* is negligible with high probability. Thus, we may focus in the analysis on $\sum_{j=1}^{k} \rho_j$. This sum can be trivially upper bounded by

$$\begin{aligned}\sum_{j=1}^{k} \rho_j &\le \sum_{j=1}^{k} \lceil \alpha \cdot z_k \rceil \\ &\le k + \sum_{j=1}^{k} \alpha \cdot z_k \\ &\le \delta^{-2} + (1+\delta^2)(1+2\delta)\delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v) \\ &\le \delta^{-2} + (1+4\delta) \cdot \delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v),\end{aligned} \tag{6}$$

where the third inequality uses $\alpha = -\ln(1-\delta) \le \delta \cdot (1+2\delta)$, $k = \lceil \log_{1-\delta}(\delta) \rceil \le \delta^{-2}$ and

$$z_j \le (1+\delta^2) \cdot \mathrm{OPT}(S_{j-1}, v) \ .$$

Following (6), we turn our attention to the expression $\delta \cdot \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v)$.

We use the next trivial bound for small values of j.

Observation 3.3. *For $j = 1, 2, \ldots, k$ it holds that* $\mathrm{OPT}(S_{j-1}, v) \le \mathrm{OPT}$.

We can use the subset-oblivious algorithm appr to obtain an additional bound on $\mathrm{OPT}(S_{j-1}, v)$. Let K and ψ be the δ^2-parameters of appr. Observe that by Definition 3.1, it holds that K and ψ depend solely on δ^2, and are independent of the instance (I, v). Without loss of generality, we assume that $\psi, K > 1$.

Lemma 3.4. *With probability at least* $1 - K \cdot \delta^2 \cdot \exp\left(-\frac{\delta^8}{\psi^2} \cdot \mathrm{OPT}\right)$, *it holds that*

$$\forall j \in \{0, 1, \ldots, k-1\} : \quad \mathrm{OPT}(S_j, v) \le \beta \cdot (1-\delta)^j \cdot \mathrm{OPT} + 2 \cdot \delta^2 \cdot \beta \cdot \mathrm{OPT} + K \ .$$

Proof. Let $\mathcal{S}$ be the set of δ^2-weight vectors of appr and (I, v). The set $\mathcal{S}$ is non-random, and is therefore $\mathcal{F}_0$-measurable. Thus, by Lemma 2.4, for every $\bar{u} \in \mathcal{S}$ it holds that

$$\begin{aligned}
& \Pr\left(\exists j \in \{0,1,\ldots,k\} : \bar{u} \cdot \mathbb{1}_{S_j} > (1-\delta)^j \cdot \|\bar{u}\| + \delta^2 \cdot \text{OPT}\right) \\
= \; & \Pr\left(\exists j \in \{0,1,\ldots,k\} : \bar{u} \cdot \mathbb{1}_{S_j} - (1-\delta)^j \cdot \bar{u} \cdot \mathbb{1}_{S_0} > \frac{\delta^2 \cdot \text{OPT}}{\mathsf{tol}(\bar{u})} \cdot \mathsf{tol}(\bar{u})\right) \\
\leq \; & \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^4 \cdot \left(\frac{\delta^2 \cdot \text{OPT}}{\mathsf{tol}(\bar{u})}\right)^2}{\text{OPT}}\right) \\
\leq \; & \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^8 \cdot \text{OPT}}{\psi^2}\right),
\end{aligned} \tag{7}$$

where the last inequality uses $\mathsf{tol}(\bar{u}) \leq \psi$. We note that the second inequality in (7) assumes $\mathsf{tol}(\bar{u}) \neq 0$, but the same outcome (i.e., the first expression is at most the last expression) can be trivially shown in case $\mathsf{tol}(\bar{u}) = 0$ (that is, $\bar{u}$ is the zero vector).

As $|\mathcal{S}| \leq K$, we can use (7) and the union bound to get

$$\Pr\left(\exists \bar{u} \in \mathcal{S}, j \in \{0,1,\ldots,k\} : \bar{u} \cdot \mathbb{1}_{S_j} > (1-\delta)^j \|\bar{u}\| + \delta^2 \text{OPT}\right) \leq K\delta^{-2} \exp\left(-\frac{2\delta^8 \text{OPT}}{\psi^2}\right) \; . \tag{8}$$

For the remainder of the proof we assume that

$$\forall j \in \{0,1,\ldots,k\}, \; \bar{u} \in \mathcal{S} : \quad \bar{u} \cdot \mathbb{1}_{S_j} \leq (1-\delta)^j \|\bar{u}\| + \delta^2 \text{OPT} \; . \tag{9}$$

By (8), this assumption holds with probability at least $1 - K \cdot \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^8 \cdot \text{OPT}}{\psi^2}\right)$.

Recall that $\text{OPT}(I, v) \geq \max_{\bar{u} \in \mathcal{S}} \|\bar{u}\|$ (Definition 3.1). Thus,

$$\forall j \in \{0,1,\ldots,k\}, \; \bar{u} \in \mathcal{S} : \quad \bar{u} \cdot \mathbb{1}_{S_j} \leq (1-\delta)^j \|\bar{u}\| + \delta^2 \text{OPT} \leq (1-\delta)^j \cdot \text{OPT} + \delta^2 \dot{\text{OPT}}, \tag{10}$$

where the first inequality is by (9). Hence, by Definition 3.1, for every $j = 1, \ldots, k$ it holds that

$$\begin{aligned}
\text{OPT}(S_{j-1}, v) \leq \; & \mathsf{appr}(I, v, S_{j-1}) \\
\leq \; & \beta \cdot \max_{\bar{u} \in \mathcal{S}} \mathbb{1}_{S_{j-1}} \cdot \bar{u} + \delta^2 \cdot \text{OPT} + K \\
\leq \; & \beta \cdot \left((1-\delta)^{j-1} \cdot \text{OPT} + \delta^2 \dot{\text{OPT}}\right) + \delta^2 \cdot \text{OPT} + K \\
\leq \; & \beta \cdot (1-\delta)^{j-1} \cdot \text{OPT} + 2 \cdot \beta\delta^2 \cdot \text{OPT} + K,
\end{aligned} \tag{11}$$

where the second inequality is by (10). Since we assumed (9) holds, (11) holds with probability at least $1 - K\delta^{-2} \exp\left(-\frac{2 \cdot \delta^8 \cdot \text{OPT}}{\psi^2}\right)$. $\square$

We note that Observation 3.3 and Lemmas 2.6 and 3.4 suffice to show that Algorithm 1 achieves an asymptotic approximation ratio arbitrarily close to $(1 + \ln \beta)$, which matches the *Round&Approx* framework. To show Algorithm 1 is strictly better we use some additional components.

We say a configuration $C \in \mathcal{C}$ has *δ-full slack* if $v_t(C) \leq 1 - \delta$ for all $t = 1, \ldots, d$. Define $\kappa(\delta) = \exp(\exp(\delta^{-3}))$.

Lemma 3.5 (Weak Structural Property). *Let $B_1, \ldots, B_s \in \mathcal{C}$ be configurations such that B_ℓ has δ-full slack for all $\ell = 1, \ldots, s$, and let $R = \bigcup_{\ell=1}^{s} B_\ell$. Then there exists a set $\mathcal{S} \subseteq \mathbb{R}^I_{\geq 0}$ such that*

- $|\mathcal{S}| \leq \kappa(\delta)$,

- $\text{supp}(\bar{u}) \subseteq R$ *for all* $\bar{u} \in \mathcal{S}$,[11]
- *and for all* $Q \subseteq R$ *and* $\gamma \in [0,1]$ *which satisfy*

$$\forall \bar{u} \in \mathcal{S}: \quad \mathbb{1}_Q \cdot \bar{u} \leq \gamma \cdot \mathbb{1}_R \cdot \bar{u} + \frac{\delta^{20}}{\kappa(\delta)} \cdot \text{OPT}(I,v) \cdot \mathsf{tol}(\bar{u}),$$

it holds that $\text{OPT}(Q,v) \leq \gamma(1+d\cdot\delta)\cdot s + \delta^{10}\cdot\text{OPT} + \kappa(\delta)$.

We refer to $\mathcal{S}$ as the *weak structure* of $B_1,\ldots,B_s$. We defer the proof of Lemma 3.5 to Section 3.2. Intuitively, Lemma 3.5 can be interpreted as follows. If R can be packed using s configurations with δ-full slack, and $Q \subseteq R$ is a random subset of R such that $\Pr(i \in Q) \leq \gamma$ then $\text{OPT}(Q,v) \lesssim \gamma s$, assuming Q satisfies some concentration bounds.

We also utilize the existence of a nearly optimal solution of (I,v) satisfying some additional properties. We say an item $i \in I$ is δ-*large* if there is a $t \in \{1,\ldots,d\}$ such that $v_t(i) \geq \delta$; otherwise, the item is *small*. Observe that these notions extend the ones given in Section 1.4 for the special case of $d = 2$. Thus, we also use L to denote the set of δ-large items in the instance (I,v).

Lemma 3.6 (Arranged solution)**.** *For any* (I,v) *there exists a solution* $A_1,\ldots,A_m$ *and sets* $W_1,\ldots,W_m \subseteq I$ *such that*

- $m \leq (1+d^2\cdot 14\cdot\delta)\cdot\text{OPT}+1$,
- $W_b \subseteq A_b \cap L$ *for* $b = 1,\ldots,m$,
- $|W_b| \leq d$ *for* $b = 1,\ldots,m$,
- *and* $A_b \setminus W_b$ *has* δ*-full slack for* $b = 1,\ldots,m$.

We refer to $A_1,\ldots,A_m$ and $W_1,\ldots,W_m$ as an *arranged solution* of (I,v). We use Lemma 3.6 as a means to utilize Lemma 3.5. The main observation is that if $Z \subseteq [m]$ is a subset of configurations in the arranged solution such that $W_b \cap S_j = \emptyset$ for every $b \in Z$, then there is a weak structure of the configurations $(A_b \cap S_j)_{b\in Z}$ which can be used to bound $\text{OPT}(S_r, v)$ for $r \geq j$.

The proof of Lemma 3.6 utilizes the following technical lemma of Bansal et al. [BEK16].

Lemma 3.7. *Let* $C \in \mathcal{C}$ *and let* $Z \subseteq [d]$ *be a set of coordinates such that* $v_t(C) > 1-\delta$ *for all* $t \in Z$ *and* $v_t(i) \leq \delta$ *for all* $i \in C$ *and* $t \in Z$*. Then there is* $Q \subseteq C$ *such that* $v_t(Q) \geq \delta$ *for all* $t \in Z$ *and* $v_t(Q) \leq 7\cdot d^2\cdot\delta$ *for all* $t \in [d]$.

Proof of Lemma 3.6. Let $A'_1,\ldots,A'_{m'}$ be an optimal solution for (I,v). That is, $m' = \text{OPT}(I,v)$.

For every $b \in [m']$ we define a set W_b as follows. Start with $W_b = \emptyset$ and while there is a coordinate $t \in [d]$ and $i \in A'_b \setminus W_b$ such that $v_t(A'_b \setminus W_b) > 1-\delta$ and $v_t(i) \geq \delta$ add the item i to W_b. Clearly, at the end of the process $|W_b| \leq d$ and $W_b \subseteq A_b \cap L$. Furthermore, let $Z_b = \{t \in \{1,\ldots,d\} \mid v_t(A'_b \setminus W_b) > 1-\delta\}$. By construction of W_b it holds that $v_t(i) \leq \delta$ for all $i \in A'_b \setminus W_b$ and $t \in Z_b$. Thus, by Lemma 3.7, for $b = 1,\ldots,m'$ there exists $Q_b \subseteq A'_b \setminus W_b$ such that $A'_b \setminus W_b \setminus Q_b$ has full slack and $v_t(Q_b) \leq 7\cdot d^2\cdot\delta$ for all $t = 1,\ldots,d$.

Define $A_b = A'_b \setminus Q_b$. By the above $A_b \setminus W_b$ has δ-full slack for $b = 1,\ldots,m'$. Let $\eta = \left\lfloor \frac{1-\delta}{7\cdot d^2 \delta} \right\rfloor$, then the union of every η of the sets among $Q_1,\ldots,Q_{m'}$ is a configuration with δ-full slack. We simply iteratively pack η of the sets $Q_1,\ldots,Q_{m'}$ into a single configuration. Thus there are configurations $A_{m'+1},\ldots,A_{m'+r}$ such that A_b is with δ-full slack for every $b = m'+1,\ldots,m'+r$, $r \leq \frac{m'}{\eta}+1$ and $A_{m'+1}\cup\ldots\cup A_{m'+r} = Q_1\cup\ldots,\cup Q_{m'}$. We define $W_{m'+1},\ldots,W_{m'+r} = \emptyset$ and $m = m'+r$.

[11] We define $\text{supp}(\bar{u}) = \{i \in I \mid \bar{u}_i > 0\}$.

Since $\delta < \frac{1}{28 \cdot d^2}$, it holds that

$$\eta \ = \ \left\lfloor \frac{1-\delta}{7 \cdot d^2 \delta} \right\rfloor \ \geq \ \frac{1-\delta}{7 \cdot d^2 \delta} - 1 = \frac{1-\delta-7 \cdot d^2 \cdot \delta}{7 \cdot d^2 \cdot \delta} \geq \frac{\frac{1}{2}}{7 \cdot d^2 \cdot \delta} = \frac{1}{14 \cdot d^2 \cdot \delta} \ .$$

Therefore, $r \leq \frac{m'}{\eta} + 1 \leq 14 \cdot d^2 \cdot \delta m' + 1 = 14 \cdot d^2 \cdot \delta \mathrm{OPT} + 1$. Hence, $m = (1 + 14 \cdot d^2 \cdot \delta) \cdot \mathrm{OPT} + 1$. □

Let $A_1, \ldots, A_m$ and $W_1, \ldots, W_m \subseteq I$ be an arranged solution of (I, v). For every $j = 0, 1, \ldots, k$ define

$$T_j = \{b \in [m] \mid W_b \cap S_j = \emptyset\} \tag{12}$$

to be the (indices of) configurations in the arranged solution such that $A_b \cap S_j$ is guaranteed to have δ-full slack. Define $j_1 = \left\lceil \frac{1}{2} \log_{1-\delta} \frac{1}{\beta} \right\rceil$.

Lemma 3.8. *With probability at least* $1 - K \cdot \kappa(\delta) \cdot \delta^{-4} \cdot \exp\left(-\frac{\delta^{50}}{\psi^2 \cdot \kappa^2(\delta)} \cdot \mathrm{OPT}\right)$*, it holds that*

$$\delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v) \ \leq \ (1+\ln\beta)\mathrm{OPT} + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2}\ln\beta\right) + 60 \cdot d^2 \beta \delta \cdot \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) \ .$$

The implication of Lemma 3.8 is that if we show that $|T_{j_1}|$ is at least a constant fraction of OPT (with high probability), then Algorithm 1 attains an asymptotic approximation ratio which is strictly better than the $(1 + \ln\beta)$ of *Round&Approx*. Indeed, such an assertion about T_{j_1} will be proved later on in Lemma 3.9. We also note that the value of j_1 was selected arbitrarily. A more refined analysis may consider $|T_j \setminus T_{j-1}|$ for *all* values of j. This concept is ingrained into our tighter analysis for the special case of 2DVP given in Section 4.1.

The proof of Lemma 3.8 partitions the sum $\delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v)$ into three parts. The first part is $\delta \sum_{j=1}^{j_1} \mathrm{OPT}(S_{j-1}, v)$, which is trivially bounded via Observation 3.3. The last part is $\delta \sum_{j=j_2+1}^{k} \mathrm{OPT}(S_{j-1}, v)$, where $j_2 = \left\lceil \log_{1-\delta} \frac{1}{\beta} \right\rceil$. Using the subset-oblivious algorithm based bound in Lemma 3.4, this sum can be bounded by roughly OPT. The (remaining) middle part, $\delta \sum_{j=j_1+1}^{j_2} \mathrm{OPT}(S_{j-1}, v)$, utilizes a weak structure of the configuration in T_{j_1} to attain a bound on $\mathrm{OPT}(S_{j-1}, v)$, which is better than the trivial bound of OPT (and also better than the bound of Lemma 3.4 which is worse for those values of j).

Proof of Lemma 3.8. By Observation 3.3,

$$\begin{aligned} \delta \sum_{r=1}^{j_1} \mathrm{OPT}(S_{r-1}, v) \leq & \ \delta \sum_{r=1}^{j_1} \mathrm{OPT} \\ = & \ \delta \cdot j_1 \cdot \mathrm{OPT} \\ \leq & \ \delta \cdot \frac{1}{2} \cdot \frac{\ln\beta}{-\ln(1-\delta)} \cdot \mathrm{OPT} + \delta \cdot \mathrm{OPT} \\ \leq & \ \frac{1}{2} (\ln\beta) \, \mathrm{OPT} + \delta \cdot \mathrm{OPT} \ . \end{aligned} \tag{13}$$

The second inequality follows from the definition of j_1, and the third inequality holds since $-\ln(1-\delta) \geq \delta$.

Assume that

$$\mathrm{OPT}(S_j, v) \leq \beta \cdot (1-\delta)^j \cdot \mathrm{OPT} + 2 \cdot \delta^2 \cdot \beta \cdot \mathrm{OPT} + K \tag{14}$$

for all $j = 0, 1, \ldots, k-1$. By Lemma 3.4, Assumption 14 holds with probability at least $1 - K \cdot \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^8 \cdot \mathrm{OPT}}{\psi^2}\right)$. Also, define $j_2 = \left\lceil \log_{1-\delta} \frac{1}{\beta} \right\rceil$; therefore,

$$
\begin{aligned}
\delta \sum_{r=j_2+1}^{k} \mathrm{OPT}(S_{r-1}, v) \leq\ & \delta \sum_{r=j_2+1}^{k} \left(\beta \cdot (1-\delta)^{r-1} \cdot \mathrm{OPT} + 2 \cdot \delta^2 \cdot \beta \cdot \mathrm{OPT} + K\right) \\
\leq\ & \beta\delta \cdot (1-\delta)^{j_2} \cdot \mathrm{OPT} \sum_{r=0}^{\infty} (1-\delta)^r + 2 \cdot k \cdot \delta \cdot \delta^2 \cdot \beta \cdot \mathrm{OPT} + \delta \cdot k \cdot K \\
\leq\ & \beta \cdot (1-\delta)^{j_2} \cdot \delta \cdot \frac{1}{1-(1-\delta)} \cdot \mathrm{OPT} + 2 \cdot \delta\beta \cdot \mathrm{OPT} + \delta^{-2} K \\
\leq\ & \beta \cdot \frac{1}{\beta} \cdot \mathrm{OPT} + 2 \cdot \delta\beta \cdot \mathrm{OPT} + \delta^{-2} \cdot K \\
\leq\ & \mathrm{OPT} + 2 \cdot \delta \cdot \beta \cdot \mathrm{OPT} + \delta^{-2} \cdot K \ .
\end{aligned}
\tag{15}
$$

The third inequality uses $k \leq \delta^{-2}$ and the forth inequality follows from $(1-\delta)^{j_2} \leq \frac{1}{\beta}$.

Let $Q^* = \bigcup_{b \in T_{j_1}} A_b \cap S_{j_1}$. That is, Q^* is the set of all items in configurations which are guaranteed to have δ-full slack in iteration j_1. Since $(A_b \cap S_{j_1})_{b \in T_{j_1}}$ is a collection of configuration with δ-full slack, by Lemma 3.5 there is a weak structure $\mathcal{S}$ of $(A_b \cap S_{j_1})_{b \in T_{j_1}}$. In particular, $\mathcal{S}$ is $\mathcal{F}_{j_1}$-measurable. Since $\mathrm{supp}(\bar{u}) \subseteq Q^*$ for all $\bar{u} \in \mathcal{S}$, it follows that $\bar{u} \cdot \mathbb{1}_{S_r} = \bar{u} \cdot \mathbb{1}_{Q^* \cap S_r}$ for all $\bar{u} \in \mathcal{S}$ and $r = j_1, j_1+1, \ldots, k$.

By Lemma 2.4, for every $\bar{u} \in \mathcal{S}$ it holds that

$$
\begin{aligned}
& \Pr\left(\exists r \in \{j_1, \ldots, k\} :\ \bar{u} \cdot \mathbb{1}_{Q^* \cap S_r} - (1-\delta)^{r-j_1} \cdot \bar{u} \cdot \mathbb{1}_{Q^*} > \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u})\right) \\
=\ & \Pr\left(\exists r \in \{j_1, \ldots, k\} :\ \bar{u} \cdot \mathbb{1}_{S_r} - (1-\delta)^{r-j_1} \cdot \bar{u} \cdot \mathbb{1}_{S_{j_1}} > \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u})\right) \\
\leq\ & \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^4 \cdot \frac{\delta^{40}}{\kappa^2(\delta)} \cdot \mathrm{OPT}^2}{\mathrm{OPT}}\right) \\
\leq\ & \delta^{-2} \cdot \exp\left(-\frac{\delta^{50}}{\kappa^2(\delta)} \cdot \mathrm{OPT}\right) \ .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \Pr\left(\forall \bar{u} \in \mathcal{S}, r \in \{j_1, \ldots, k\} :\ \bar{u} \cdot \mathbb{1}_{S_r \cap Q^*} \leq (1-\delta)^{r-j_1} \cdot \bar{u} \cdot \mathbb{1}_{Q^*} + \delta^{20} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u})\right) \\
\geq\ & 1 - |\mathcal{S}| \cdot \delta^{-2} \cdot \exp\left(-\frac{\delta^{50}}{\kappa^2(\delta)} \cdot \mathrm{OPT}\right) \\
\geq\ & 1 - \kappa(\delta) \cdot \delta^{-2} \cdot \exp\left(-\frac{\delta^{50}}{\kappa^2(\delta)} \cdot \mathrm{OPT}\right) \ .
\end{aligned}
\tag{16}
$$

For the remainder of the proof we assume that

$$
\forall \bar{u} \in \mathcal{S}, r \in \{j_1, \ldots, k\} :\ \bar{u} \cdot \mathbb{1}_{S_r \cap Q^*} \leq (1-\delta)^{r-j_1} \cdot \bar{u} \cdot \mathbb{1}_{Q^*} + \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u}). \tag{17}
$$

By (16), this assumption holds with probability at least $1 - \kappa(\delta) \cdot \delta^{-2} \cdot \exp\left(-\frac{\delta^{50}}{\kappa^2(\delta)} \cdot \mathrm{OPT}\right)$.

By (17) it holds that

$$\mathrm{OPT}(S_r \cap Q^*, v) \leq (1-\delta)^{r-j_1} \cdot (1+d\cdot\delta)|T_{j_1}| + \delta^{10} \cdot \mathrm{OPT} + \kappa(\delta) \tag{18}$$

for all $r = j_1, j_1+1, \ldots k$. It trivially holds that

$$\delta \sum_{r=j_1+1}^{j_2} \mathrm{OPT}(S_{r-1}, v) = \delta \sum_{r=j_1+1}^{j_2} \mathrm{OPT}(S_{r-1} \cap Q^*, v) + \delta \sum_{r=j_1+1}^{j_2} \mathrm{OPT}(S_{r-1} \setminus Q^*, v) \ . \tag{19}$$

By (18) we have

$$\begin{aligned}
\delta \sum_{r=j_1+1}^{j_2} \mathrm{OPT}(S_{r-1} \cap Q^*, v) \leq&\ \delta \sum_{r=j_1+1}^{j_2} \left((1-\delta)^{r-1-j_1} \cdot (1+d\cdot\delta)|T_{j_1}| + \delta^{10}\cdot\mathrm{OPT} + \kappa(\delta)\right) \\
\leq&\ \delta(1+d\cdot\delta)\cdot|T_{j_1}| \sum_{r=j_1+1}^{j_2} (1-\delta)^{r-1-j_1} + k\cdot\delta^{11}\cdot\mathrm{OPT} + \delta\cdot k\cdot\kappa(\delta) \\
\leq&\ (1+d\cdot\delta)\cdot|T_{j_1}|\cdot\delta\cdot\frac{1-(1-\delta)^{j_2-1-j_1+1}}{1-(1-\delta)} + \delta^9\cdot\mathrm{OPT} + \delta^{-1}\cdot\kappa(\delta) \\
\leq&\ (1+d\cdot\delta)\cdot|T_{j_1}|\left(1 - \frac{1}{\sqrt{\beta}}(1-\delta)\right) + \delta^9\cdot\mathrm{OPT} + \delta^{-1}\cdot\kappa(\delta) \\
\leq&\ |T_{j_1}|\cdot\left(1-\frac{1}{\sqrt{\beta}}\right) + 10\cdot d\cdot\delta\cdot\mathrm{OPT} + \delta^{-1}\cdot\kappa(\delta) \ .
\end{aligned} \tag{20}$$

The second inequality holds as $j_2 - j_1 \leq k$. The third inequality uses $k \leq \delta^{-2}$. The forth inequality holds, as

$$j_2 - j_1 \leq \log_{1-\delta}\frac{1}{\beta} + 1 - \frac{1}{2}\cdot\log_{1-\delta}\frac{1}{\beta} = \frac{1}{2}\cdot\log_{1-\delta}\frac{1}{\beta} + 1, \tag{21}$$

thus $(1-\delta)^{j_2-j_1} \geq \frac{1}{\sqrt{\beta}}\cdot(1-\delta)$. The fifth inequality holds, as $|T_{j_1}| \leq m \leq (1+14\cdot d^2\cdot\delta)\mathrm{OPT}+1 \leq 2\cdot\mathrm{OPT}+1$.

It trivially holds that $\mathrm{OPT}(S_{r-1}\setminus Q^*, v) \leq m - |T_j|$ for $r = j_1+1,\ldots,j_2$ via the configurations $(A_b \cap S_{r-1})_{b\in\{1,\ldots,m\}\setminus T_j}$. Therefore,

$$\begin{aligned}
\delta \sum_{r=j_1+1}^{j_2} \mathrm{OPT}(S_{r-1}\setminus Q^*, v) \leq&\ \delta(j_2-j_1)(m-|T_{j_1}|) \\
\leq&\ \delta\left(\frac{1}{2}\cdot\log_{1-\delta}\frac{1}{\beta}+1\right)\cdot(m-|T_{j_1}|) \\
=&\ \delta\cdot\frac{1}{2}\cdot\frac{\ln\beta}{-\ln(1-\delta)}\cdot(m-|T_{j_1}|) + \delta(m-|T_{j_1}|) \\
\leq&\ \frac{1}{2}(\ln\beta)(m-|T_{j_1}|) + \delta m \\
\leq&\ \frac{1}{2}(\ln\beta)(m-|T_{j_1}|) + 2\cdot\delta\mathrm{OPT} \ .
\end{aligned} \tag{22}$$

The second inequality follows from (21). The third inequality holds, as $-\ln(1-\delta) \geq \delta$.

By (19), (20) and (22) we have

$$
\begin{aligned}
&\delta \sum_{r=j_1+1}^{j_2} \mathrm{OPT}(S_{r-1}, v) \\
&\quad \leq \ |T_{j_1}| \left(1 - \frac{1}{\sqrt{\beta}}\right) + 10 \cdot d \cdot \delta \mathrm{OPT} + \delta^{-1} \cdot \kappa(\delta) + \frac{1}{2} (\ln \beta) (m - |T_{j_1}|) + 2 \cdot \delta \mathrm{OPT} \\
&\quad \leq \ \frac{m}{2} \ln \beta + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2} \ln \beta\right) + 20 \cdot d \cdot \delta \cdot \mathrm{OPT} + \delta^{-1} \cdot \kappa(\delta) \\
&\quad \leq \ \frac{(1 + 14 \cdot d^2 \delta) \cdot \mathrm{OPT} + 1}{2} \ln \beta + |T_{j_1}| \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2} \ln \beta\right) + 20 \cdot d\delta \mathrm{OPT} + \delta^{-1} \kappa(\delta) \\
&\quad \leq \ \frac{\mathrm{OPT}}{2} \ln \beta + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2} \ln \beta\right) + 50 \cdot d^2 \cdot \beta \cdot \delta \cdot \mathrm{OPT} + \delta^{-1} \cdot \kappa(\delta) + \frac{\ln \beta}{2} \ .
\end{aligned} \tag{23}
$$

By (13), (15), and (23) we have

$$
\begin{aligned}
&\delta \sum_{r=1}^{k} \mathrm{OPT}(S_{r-1}, v) \\
&\quad \leq \ \frac{1}{2} (\ln \beta)\, \mathrm{OPT} + \delta \mathrm{OPT} \\
&\qquad\quad + \frac{\mathrm{OPT}}{2} \ln \beta + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2} \ln \beta\right) + 50 \cdot d^2 \beta \delta \cdot \mathrm{OPT} + \delta^{-1} \cdot \kappa(\delta) + \frac{\ln \beta}{2} \\
&\qquad\quad + \mathrm{OPT} + 2 \cdot \delta \cdot \beta \cdot \mathrm{OPT} + \delta^{-2} \cdot K \\
&\quad \leq \ (1 + \ln \beta) \mathrm{OPT} + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2} \ln \beta\right) + 60 \cdot d^2 \beta \delta \cdot \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) \ .
\end{aligned}
$$

As we assumed that (14) and (17) hold, the statement holds with probability

$$
\begin{aligned}
&1 - \kappa(\delta) \cdot \delta^{-2} \cdot \exp\left(-\frac{\delta^{50}}{\kappa^2(\delta)} \cdot \mathrm{OPT}\right) - K \cdot \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^8 \cdot \mathrm{OPT}}{\psi^2}\right) \\
\geq \ &1 - K \cdot \kappa(\delta) \cdot \delta^{-4} \cdot \exp\left(-\frac{\delta^{50}}{\psi^2 \cdot \kappa^2(\delta)} \cdot \mathrm{OPT}\right) \ . *
\end{aligned}
$$

□

To attain the statement of Theorem 3.2, we show that $|T_{j_1}|$ is at least a constant fraction of OPT (with high probability), and combine this result with Lemma 3.8.

Lemma 3.9. *With probability at least* $1 - \delta^{-2} \cdot \exp(-\delta^{50} \cdot \mathrm{OPT})$*, it holds that*

$$
|T_{j_1}| \geq \left(1 - \beta^{-\frac{1}{2d}}\right)^d \cdot m - 4 \cdot d \cdot \delta \cdot \mathrm{OPT} \ .
$$

Proof. For $j = 0, 1, \ldots, k$ and $\ell = 0, 1, \ldots, d$ define $V_{j,\ell} = \{b \in [m] \mid |W_b \cap S_j| = \ell\}$ and $V_{j,\leq \ell} = \bigcup_{h=0}^{\ell} V_{j,\ell}$. Since $A_b \setminus W_b$ is guaranteed to have δ-full slack, the set $V_{j,\ell}$ ($V_{j,\leq \ell}$) can be intuitively interpreted as (the indices of) the set of configurations among $A_1 \cap S_j, \ldots, A_m \cap S_j$ which have δ-full slack if (at most) ℓ specific large items are removed from them. Since $S_0 \supseteq S_1 \supseteq \ldots \supseteq S_k$ it holds that $V_{0,\leq \ell} \subseteq V_{1,\leq \ell} \subseteq \ldots \subseteq V_{k,\leq \ell}$. Observe that $V_{j,\ell}$ is $\mathcal{F}_j$-measurable and $T_j = V_{j,0} = V_{j,\leq 0}$.

Observe that for every $b = 1, 2, \ldots, m$ and $\ell = 0, 1, \ldots, d$ it holds that $\{j \mid b \in V_{j,\ell}\}$ is a set of consecutive integers. That is, b belong to $V_{j,\ell}$ from some iteration r_1 up to some iteration r_2. The next claim essentially states that the difference $r_2 - r_1$ is not expected to be too large.

Claim 3.10. *Let $j \in \{0, 1 \ldots, k-1\}$, $\ell \in \{1, \ldots, d\}$ and let $Z \subseteq V_{j,\ell}$ be an $\mathcal{F}_j$-measurable subset. Then it holds that*

$$\mathbb{E}\left[\, |Z \cap V_{j,\ell+1}| \mid \mathcal{F}_j\right] \leq (1-\delta) \cdot |Z| \ .$$

Proof. For $b = 1, \ldots, m$ let i_b be an arbitrary item in $W_b \cap S_j$ (or an arbitrary item in I in case $W_b \cap S_j = \emptyset$). In particular, i_b is an $\mathcal{F}_j$-measurable random variable. For $b = 1, \ldots, m$ it holds that

$$\begin{aligned}
\Pr\left(b \in Z \cap V_{j+1,\ell} \mid \mathcal{F}_j\right) &= \Pr\left(b \in Z \text{ and } W_b \cap S_j \subseteq S_{j+1} \mid \mathcal{F}_j\right) \\
&\leq \Pr\left(b \in Z \text{ and } i_b \in S_{j+1} \mid \mathcal{F}_j\right) \\
&\leq \mathbb{1}_{b \in Z} \cdot (1-\delta) \cdot \mathbb{1}_{i_b \in S_j} \\
&= (1-\delta) \cdot \mathbb{1}_{b \in Z} \ .
\end{aligned}$$

The second inequality follows from Lemma 2.1. That last equality holds since if $b \in Z \subseteq V_{j,\ell}$ then $i_b \in S_j$ as $\ell \neq 0$. Thus,

$$\mathbb{E}\left[\, |Z \cap V_{j,\ell+1}| \mid \mathcal{F}_j\right] = \sum_{b \in [m]} \Pr\left(b \in Z \cap V_{j+1,\ell} \mid \mathcal{F}_j\right) \leq \sum_{b \in [m]} (1-\delta) \cdot \mathbb{1}_{b \in Z} = (1-\delta)|Z| \ .$$

◇

We use Lemma 2.3 to show that $|Z \cap V_{j,\ell+1}|$ cannot be significantly larger than the bound on its expectation as stated in Claim 3.10.

Claim 3.11. *Let $j \in \{0, 1 \ldots, k-1\}$, $\ell \in \{1, \ldots, d\}$ and let $Z \subseteq V_{j,\ell}$ be an $\mathcal{F}_j$-measurable subset. Then $|Z \cap V_{j,\ell+1}| \leq (1-\delta) \cdot |Z| + \delta^{20} \cdot \mathrm{OPT}$ with probability at least $1 - \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$.*

Proof. For every $S \subseteq I$, $\rho \in [\mathrm{OPT}]$ and $X \subseteq [m]$ define a function $f_{S,\rho,X} : \mathcal{C}^{\mathrm{OPT}} \to \mathbb{R}$ by

$$f_{S,\rho,X}(C_1, \ldots, C_{\mathrm{OPT}}) = \sum_{b \in X} \mathbb{1}_{W_b \cap S \cap \left(\bigcup_{s=1}^{\rho} C_s\right) = \emptyset} \ .$$

Observe that

$$f_{S_j,\rho_{j+1},Z}(C_1^{j+1}, \ldots, C_{\mathrm{OPT}}^{j+1}) \ = \ \sum_{b \in Z} \mathbb{1}_{W_b \cap S_j \cap \left(\bigcup_{s=1}^{\rho_{j+1}} C_s^{j+1}\right) = \emptyset} \ = \ \sum_{b \in Z} \mathbb{1}_{Z \in V_{j+1,\ell}} \ = \ |Z \cap V_{j+1,\ell}| \ .$$

Moreover, as S_j, ρ_{j+1} and Z are $\mathcal{F}_j$ measurable it follows that $f_{S_j,\rho_{j+1},Z}$ is $\mathcal{F}_j$-measurable as well (note that ρ_{j+1} is determined before $C^{j_1}{}_{,1}, \ldots, C_{\rho_{j+1}}^{j+1}$ are sampled in Line 3 of Algorithm 1).

Define $D = \{f_{S,\rho,X} \mid S \subseteq I,\ \rho \in [\mathrm{OPT}],\ X \subseteq [m]\}$. It follows that D is a finite set. In order to use Lemma 2.3 we need to show that the functions in D are of bounded difference.

Let $f_{S,\rho,X} \in D$, $(C_1, \ldots, C_{\mathrm{OPT}})$, $(C'_1, \ldots, C'_{\mathrm{OPT}}) \in \mathcal{C}^{\mathrm{OPT}}$ and $r \in [\mathrm{OPT}]$ such that $C_s = C'_s$ for $s = 1, \ldots, r-1, r+1, \ldots, \mathrm{OPT}$ (i.e., $(C_1, \ldots, C_{\mathrm{OPT}})$ and $(C'_1, \ldots, C'_{\mathrm{OPT}})$ are identical in all coordinates expect the r-th). If $r > \rho$ then

$$\left|f_{S,\rho,X}(C_1, \ldots, C_{\mathrm{OPT}}) - f_{S,\rho,X}(C'_1, \ldots, C'_{\mathrm{OPT}})\right| = 0 \ .$$

Otherwise,

$$\begin{aligned}
\left|f_{S,\rho,X}(C_1, \ldots, C_{\mathrm{OPT}}) - f_{S,\rho,X}(C'_1, \ldots, C'_{\mathrm{OPT}})\right| &= \left|\sum_{b \in X} \mathbb{1}_{W_b \cap S \cap \left(\bigcup_{s=1}^{\rho} C_s\right) = \emptyset} - \sum_{b \in X} \mathbb{1}_{W_b \cap S \cap \left(\bigcup_{s=1}^{\rho} C'_s\right) = \emptyset}\right| \\
&\leq \sum_{b \in X} \mathbb{1}_{W_b \cap S \cap C_r \neq \emptyset} + \sum_{b \in X} \mathbb{1}_{W_b \cap S \cap C'_r \neq \emptyset} \\
&\leq 2 \cdot d \cdot \delta^{-1} \ .
\end{aligned}$$

The last inequality holds, since the sets $W_1, \ldots, W_m$ are pairwise disjoint and only contain large items, and furthermore, a configuration $C \in \mathcal{C}$ may contain at most $d \cdot \delta^{-1}$ large items. Thus, $f_{S,\rho,X}$ is of $(2 \cdot d \cdot \delta^{-1})$-bounded difference.

By Claim 3.10 and Lemma 2.3 we have

$$\begin{aligned}
&\Pr\left(|Z \cap V_{j,\ell+1}| \;>\; (1-\delta)|Z| + \delta^{20} \cdot \mathrm{OPT}\right) \\
&\quad\le\; \Pr\Big(|Z \cap V_{j,\ell+1}| - \mathbb{E}\big[|Z \cap V_{j,\ell+1}| \;\big|\; \mathcal{F}_j\big] \;>\; \delta^{20} \cdot \mathrm{OPT}\Big) \\
&\quad\le\; \Pr\left(f_{S_j,\rho_{j+1},Z}(C_1^{j+1}, \ldots, C_{\mathrm{OPT}}^{j+1}) - \mathbb{E}\left[f_{S_j,\rho_{j+1},Z}(C_1^{j+1}, \ldots, C_{\mathrm{OPT}}^{j+1}) \;\middle|\; \mathcal{F}_j\right] \;>\; \delta^{20} \cdot \mathrm{OPT}\right) \\
&\quad\le\; \exp\left(-\frac{2 \cdot \delta^{40} \cdot \mathrm{OPT}^2}{\mathrm{OPT} \cdot 4 \cdot d^2 \cdot \delta^{-2}}\right) \;\le \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right) \;.
\end{aligned}$$

The last inequality holds as $\frac{1}{d^2} \ge 28\delta \ge \delta$. $\diamond$

Define $\eta = \left\lfloor \frac{1}{2d} \cdot \log_{1-\delta} \frac{1}{\beta} \right\rfloor$. We use Claim 3.11 to prove the following.

Claim 3.12. *Let* $\ell \in \{0, 1, \ldots, d-1\}$. *Then*

$$\left|V_{\ell\cdot\eta,d-\ell} \cap V_{(\ell+1)\cdot\eta,d-\ell}\right| \le \beta^{-\frac{1}{2d}}(1+2\delta) \cdot |V_{\ell\cdot\eta,d-\ell}| + \eta \cdot \delta^{20} \cdot \mathrm{OPT}$$

with probability at least $1 - \eta \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$.

Proof. We use induction on $j = 0, 1, \ldots, \eta$ to show that

$$|V_{\ell\cdot\eta,d-\ell} \cap V_{\ell\cdot\eta+j,d-\ell}| \le (1-\delta)^j \cdot |V_{\ell\cdot\eta,d-\ell}| + j \cdot \delta^{20} \cdot \mathrm{OPT}$$

with probability at least $1 - j \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$.

Base case: For $j = 0$, it holds that $|V_{\ell\cdot\eta,d-\ell} \cap V_{\ell\cdot\eta+j,d-\ell}| = |V_{\ell\cdot\eta,d-\ell}|$ with probability 1.

Induction Step: Assume the induction hypothesis holds for some $j \ge 0$. Define $Z = V_{\ell\cdot\eta,d-\ell} \cap V_{\ell\cdot\eta+j,d-\ell}$, and observe that Z is $\mathcal{F}_{\ell\cdot\eta+j}$-mesuarable. By the induction hypothesis and Claim 3.10, it holds that

$$\begin{array}{rrcl}
 & |Z| \;=\; |V_{\ell\cdot\eta,d-\ell} \cap V_{\ell\cdot\eta+j,d-\ell}| & \le & (1-\delta)^j \cdot |V_{\ell\cdot\eta,d-\ell}| + j \cdot \delta^{20} \cdot \mathrm{OPT} \\
\text{and} & |Z \cap V_{\eta\cdot\ell+j+1,d-\ell}| & \le & (1-\delta) \cdot |Z| + \delta^{20} \cdot \mathrm{OPT}
\end{array} \tag{24}$$

with probability at least $1 - (j+1) \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$. Furthermore, if (24) holds, then

$$\begin{aligned}
|V_{\ell\cdot\eta,d-\ell} \cap V_{\ell\cdot\eta+j+1,d-\ell}| \;&=\; |Z \cap V_{\ell\cdot\eta+j+1,d-\ell}| \\
&\le\; (1-\delta) \cdot |Z| + \delta^{20} \cdot \mathrm{OPT} \\
&\le\; (1-\delta)^{j+1} \cdot |V_{\eta\cdot\ell,d-\ell}| + (j+1) \cdot \delta^{20} \cdot \mathrm{OPT},
\end{aligned}$$

where the first equality holds since for all $b \in V_{\ell\cdot\eta,d-\ell} \cap V_{\ell\cdot\eta+j+1,d-\ell}$ it also must hold that $b \in V_{\ell\cdot\eta+j,d-\ell}$. This completes the induction step.

Therefore, using the definition of η,

$$\begin{aligned}
\left|V_{\ell\cdot\eta,d-\ell} \cap V_{(\ell+1)\cdot\eta,d-\ell}\right| \;&\le\; (1-\delta)^\eta \cdot |V_{\ell\cdot\eta,d-\ell}| + \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\
&\le\; \frac{\beta^{-\frac{1}{2d}}}{1-\delta} \cdot |V_{\ell\cdot\eta,d-\ell}| + \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\
&\le\; \beta^{-\frac{1}{2d}} \cdot (1+2\delta) \cdot |V_{\ell\cdot\eta,d-\ell}| + \eta \cdot \delta^{20} \cdot \mathrm{OPT}
\end{aligned}$$

with probability at least $1 - \eta \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$. $\diamond$

Using Claim 3.12 and a simple induction, we attain the following.

Claim 3.13. *Let $\ell \in \{0, 1, \ldots, d\}$. Then $|V_{\ell\cdot\eta,\le d-\ell}| \ge \left(1 - \beta^{-\frac{1}{2d}}(1+2\delta)\right)^{\ell} \cdot m - \ell \cdot \eta \cdot \delta^{20} \cdot \mathrm{OPT}$ with probability at least $1 - \ell \cdot \eta \cdot \exp(-\delta^{50} \cdot \mathrm{OPT})$.*

Proof. We prove the claim by induction over ℓ.
Base Case: For $\ell = 0$ it holds that

$$|V_{0,\le d}| = |[m]| = \left(1 - \beta^{-\frac{1}{2d}} \cdot (1+2\delta)\right)^{0} \cdot m - 0 \cdot \eta \cdot \delta^{10} \cdot \mathrm{OPT} \ .$$

Induction Step: Assume the claim holds for $\ell < d$. Then, by the induction hypothesis and Claim 3.12 it holds that, with probability at least $1 - (\ell+1) \cdot \eta \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$,

$$\begin{aligned} |V_{\ell\cdot\eta,\le d-\ell}| \quad &\ge \left(1 - \beta^{-\frac{1}{2d}}(1+2\delta)\right)^{\ell} \cdot m - \ell \cdot \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\ \text{and } \left|V_{\ell\cdot\eta,d-\ell} \cap V_{(\ell+1)\cdot\eta,d-\ell}\right| \quad &\le \beta^{-\frac{1}{2d}} \cdot (1+2\delta)|V_{\ell\cdot\eta,d-\ell}| + \eta \cdot \delta^{20} \cdot \mathrm{OPT} \ . \end{aligned} \tag{25}$$

Assuming (25) holds, we have

$$\begin{aligned} \left|V_{(\ell+1)\eta,\le d-\ell-1}\right| \ &\ge \ |V_{\ell\eta,\le d-\ell-1}| + \left|V_{\ell\eta,d-\ell} \setminus V_{(\ell+1)\eta,d-\ell}\right| \\ \left|V_{(\ell+1)\eta,\le d-\ell-1}\right| \ &\ge \ |V_{\ell\eta,\le d-\ell-1}| + \left|V_{\ell\eta,d-\ell} \setminus V_{(\ell+1)\eta,d-\ell}\right| \\ &= \ |V_{\ell\eta,\le d-\ell-1}| + |V_{\ell\eta,d-\ell}| - \left|V_{\ell\eta,d-\ell} \cap V_{(\ell+1)\eta,d-\ell}\right| \\ &\ge \ |V_{\ell\eta,\le d-\ell-1}| + |V_{\ell\eta,d-\ell}| - \left(\beta^{-\frac{1}{2d}} \cdot (1+2\delta)|V_{\ell\cdot\eta,d-\ell}| + \eta \cdot \delta^{20} \cdot \mathrm{OPT}\right) \\ &= \ |V_{\ell\eta,\le d-\ell-1}| + \left(1 - \beta^{-\frac{1}{2d}} \cdot (1+2\delta)\right) \cdot |V_{\ell\cdot\eta,d-\ell}| - \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\ &\ge \ \left(1 - \beta^{-\frac{1}{2d}} \cdot (1+2\delta)\right) \cdot |V_{\ell\cdot\eta,\le d-\ell}| - \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\ &\ge \left(1 - \beta^{-\frac{1}{2d}} \cdot (1+2\delta)\right) \cdot \left(\left(1 - \beta^{-\frac{1}{2d}}(1+2\delta)\right)^{\ell} \cdot m - \ell \cdot \eta \cdot \delta^{20} \cdot \mathrm{OPT}\right) - \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\ &\ge \ \left(1 - \beta^{-\frac{1}{2d}}(1+2\delta)\right)^{\ell+1} \cdot m - (\ell+1) \cdot \eta \cdot \delta^{20} \cdot \mathrm{OPT}, \end{aligned}$$

which completes the induction step. $\diamond$

By Claim 3.13 it follows that with probability at least $1 - d \cdot \eta \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right) \ge 1 - \delta^2 \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$ it holds that

$$\begin{aligned} |V_{j_1,0}| \ &\ge \ |V_{\eta\cdot d,\le 0}| \\ &\ge \ \left(1 - \beta^{-\frac{1}{2d}}(1+2\delta)\right)^{d} \cdot m - d \cdot \eta \cdot \delta^{20} \cdot \mathrm{OPT} \\ &\ge \ \left(1 - \beta^{-\frac{1}{2d}}\right)^{d} \cdot m - 2d \cdot \delta \cdot m + \cdot\delta^{18} \cdot \mathrm{OPT} \\ &\ge \ \left(1 - \beta^{-\frac{1}{2d}}\right)^{d} \cdot m - 4d \cdot \delta \cdot \mathrm{OPT} \ . \end{aligned}$$

The first inequality holds since $\eta \cdot d = d \cdot \left\lfloor \frac{1}{2d} \cdot \log_{1-\delta} \frac{1}{\beta} \right\rfloor \le \frac{1}{2} \cdot \log_{1-\delta} \frac{1}{\beta} \le j_1$. The third inequality holds since $\left(1 - \beta^{-\frac{1}{2d}}(1+2\delta)\right)^{d} \ge \left(1 - \beta^{-\frac{1}{2d}}\right)^{d} - 2\delta \cdot d$ and $d \cdot \eta \le k \le \delta^{-2}$. The last inequality holds as $m \le 2 \cdot \mathrm{OPT}$. $\square$

To complete the proof of Theorem 3.2 we only need to combine the results of Lemmas 2.6, 3.8 and 3.9. Assume the inequalities

$$
\begin{aligned}
\rho^* &\leq 8 \cdot d \cdot \delta \cdot \mathrm{OPT} + 1 \\
\delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v) &\leq (1 + \ln \beta)\mathrm{OPT} + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2}\ln\beta\right) \\
&\qquad + 60 \cdot d^2 \beta\delta \cdot \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) \\
|T_{j_1}| &\geq \left(1 - \beta^{-\frac{1}{2d}}\right)^d \cdot m - 4 \cdot d \cdot \delta \cdot \mathrm{OPT}
\end{aligned}
\tag{26}
$$

hold. By Lemmas 2.6, 3.8 and 3.9, these inequalities hold with probability at least

$$
\begin{aligned}
& 1 - \delta^{-2} \cdot \exp\left(-\delta^7 \cdot \mathrm{OPT}\right) - K \cdot \kappa(\delta) \cdot \delta^{-4} \cdot \exp\left(-\frac{\delta^{50}}{\psi^2 \cdot \kappa^2(\delta)} \cdot \mathrm{OPT}\right) - \delta^{-2} \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right) \\
\geq\ & 1 - K \cdot \delta^{-5} \cdot \kappa(\delta) \cdot \exp\left(-\frac{\delta^{50}}{\psi^2 \cdot \kappa^2(\delta)} \cdot \mathrm{OPT}\right) .
\end{aligned}
$$

Thus, if OPT is sufficiently large then (26) occurs with probability at least $\frac{1}{2}$. Furthermore, in this case it also holds that

$$
\begin{aligned}
& \delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v) \\
&\quad \leq (1 + \ln \beta)\mathrm{OPT} + |T_{j_1}| \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2}\ln\beta\right) + 60 \cdot d^2\beta\delta \cdot \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) \\
&\quad \leq (1 + \ln \beta)\mathrm{OPT} + \left(\left(1 - \beta^{-\frac{1}{2d}}\right)^d \cdot m - 4 \cdot d \cdot \delta \cdot \mathrm{OPT}\right) \cdot \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2}\ln\beta\right) \\
&\qquad\quad + 60 \cdot d^2\beta\delta \cdot \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) \\
&\quad \leq (1 + \ln \beta)\mathrm{OPT} - \chi(\beta, d) \cdot m + 90 \cdot d^2 \cdot \delta \cdot \beta \cdot \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) \\
&\quad \leq \left(1 + \ln \beta - \chi(\beta, d) + 90 \cdot d^2 \cdot \delta \cdot \beta\right) \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta) .
\end{aligned}
\tag{27}
$$

The second and third inequalities hold as $-\beta \leq \left(1 - \frac{1}{\sqrt{\beta}} - \frac{1}{2}\ln\beta\right) \leq 0$. The third inequality uses the definition of $\chi(\beta, d)$ as given in the statement of Theorem 1.5. The forth inequality holds as $\chi(\beta, d) \geq 0$ and $m \geq \mathrm{OPT}$.

By (6), (26) and (27), the size of the solution returned by Algorithm 1 is

$$
\begin{aligned}
\sum_{j=1}^{k} \rho_j + \rho^* &\leq \delta^{-2} + (1 + 4\delta) \cdot \delta \sum_{j=1}^{k} \mathrm{OPT}(S_{j-1}, v) + 8 \cdot d \cdot \delta \cdot \mathrm{OPT} + 1 \\
&\leq \delta^{-2} + (1 + 4\delta) \cdot \left(\left(1 + \ln \beta - \chi(\beta, d) + 90 \cdot d^2 \cdot \delta \cdot \beta\right) \mathrm{OPT} + \delta^{-3} K \cdot \beta \cdot \kappa(\delta)\right) + 8 \cdot d \cdot \delta \cdot \mathrm{OPT} \\
&\leq \left(1 + \ln \beta - \chi(\beta, d) + 200 \cdot d^2 \cdot \delta \cdot \beta\right) \mathrm{OPT} + \delta^{-5} K \cdot \beta \cdot \kappa(\delta) .
\end{aligned}
$$

That is, the algorithm is a randomized asymptotic $(1 + \ln\beta - \chi(\beta, d) + 200 \cdot d^2 \cdot \delta \cdot \beta)$-approximation algorithm for dVBP. □

3.2 The Weak Structural Property

In this section we prove Lemma 3.5. The lemma relies on an implicit rounding of the large items volumes to multiplicities of $\frac{\delta^2}{2d}$. While the volume of the items is rounded up, the slack of the configurations $B_1, \ldots, B_s$ ensures that these remain feasible configurations with respect to the rounded weight. Subsequently, the proof of the lemma views items of the same rounded volume as interchangeable, which is key in attaing the bound on $\mathrm{OPT}(Q, v)$ as stated in Lemma 3.5.

The lemma is utilizes some ideas from Bansal et al. [BEK16]. However, the rounding procedure in their work only requires each of the configurations $B_1, \ldots, B_s$ to have slack in $d-1$ dimension, and combines a shifting argument as part of the rounding. As mentioned in the introduction (see also Appendix A), the approach taken by Bansal et al. [BEK16] has a flaw in the analysis, and hence cannot be used. Requiring the configurations to have δ-full slack is a simple way to work around the flaw. When possible, the notations used in both lemmas are kept similar.

Proof of Lemma 3.5. We assume that $d \in \mathbb{N}_{>0}$ and $\delta \in (0, 0.1)$. Throughout the proof, consider an instance (I, v) of dVBP. Furthermore, we assume $\delta \leq \frac{1}{d^2}$ and $\delta^{-1} \in \mathbb{N}$. As in the statement of Lemma 3.5, let $B_1, \ldots, B_s \in \mathcal{C}$ be collection of configurations with δ-full slack, and define $R = B_1 \cup B_2 \cup \ldots \cup B_s$.

Recall that L is the set of large items of the instance (I, v). Set $h = \delta^{-2}$ and $\mathcal{G} = \{1, \ldots, 2 \cdot d \cdot h\}^d$. For every $\bar{a} \in \mathcal{G}$ define

$$I_{\bar{a}} = \left\{ i \in L \cap R \;\middle|\; \forall r \in [d] : \quad \frac{\delta^2}{2 \cdot d} \cdot (\bar{a}_r - 1) \;<\; v_r(i) \;\leq\; \frac{\delta^2}{2 \cdot d} \cdot \bar{a}_r \right\} \; . \tag{28}$$

Also, define the *rounded volume* of $\bar{a} \in \mathcal{G}$ by

$$\tilde{v}(\bar{a}) = \frac{\delta^2}{2 \cdot d} \cdot \bar{a} \; . \tag{29}$$

Implicitly, we round the volume of all items in $I_{\bar{a}}$ to $\tilde{v}(\bar{a})$. Since $v(i) \in (0, 1]^d$ for every $i \in I$, it follows that $\bigcup_{\bar{a} \in \mathcal{G}} I_{\bar{a}} = R \cap L$.

The *type* of a configuration $C \in \mathcal{C}$, denoted $\mathsf{T}(C)$, is the vector $\bar{t} \in \mathbb{N}^{\mathcal{G}}$ defined by $\bar{t}_{\bar{a}} = |I_{\bar{a}} \cap C|$ for every $\bar{a} \in \mathcal{G}$. That is, $\bar{t}_{\bar{a}}$ is the number of items from $I_{\bar{a}}$ in the configuration $C \in \mathcal{C}$. Define $\mathcal{T} = \{\mathsf{T}(B_\ell) \mid \ell = 1, 2, \ldots, s\}$ to be the set of all types of configurations in $B_1, \ldots, B_s$. As a configuration C may contain up to $d \cdot \delta^{-1}$ large items, it follows that

$$\begin{aligned} |\mathcal{T}| \leq \;& \left(d \cdot \delta^{-1}\right)^{|\mathcal{G}|} \\ \leq \;& \left(d \cdot \delta^{-1}\right)^{(2 \cdot d \cdot h)^d} \\ = \;& \exp\left(\left(2 \cdot d \cdot \delta^{-2}\right)^d \ln\left(d \cdot \delta^{-1}\right)\right) \\ \leq \;& \exp\left(\delta^{-4 \cdot \delta^{-1}} \ln(\delta^{-2})\right) \qquad\qquad s \\ \leq \;& \exp\left(\delta^{-5 \cdot \delta^{-1}}\right) \\ \leq \;& \frac{\kappa(\delta)}{3}, \leq \exp\left(\left(d^2 \cdot \delta^{-2}\right)^{d+1}\right). \end{aligned} \tag{30}$$

where the third inequality holds as $d^2 \leq \delta^{-1}$. Similarly to (29), we define the *rounded volume* of $\bar{t} \in \mathcal{T}$ by

$$\tilde{v}(\bar{t}) = \sum_{\bar{a} \in \mathcal{G}} \bar{t}_{\bar{a}} \cdot \tilde{v}(\bar{a}) \; . \tag{31}$$

For every $\bar{t} \in \mathcal{T}$ define

$$L_{\bar{t}} = \bigcup_{\ell \in [s] \text{ s.t. } \mathsf{T}(B_\ell) = \bar{t}} B_\ell \cap L \qquad \text{and} \qquad S_{\bar{t}} = \bigcup_{\ell \in [s] \text{ s.t. } \mathsf{T}(B_\ell) = \bar{t}} B_\ell \setminus L,$$

as the set of large items and the set of small items in configuration of type $\bar{t}$ among $B_1, \ldots, B_s$, respectively. Also, for every $r = 1, \ldots, d$ define $\bar{v}^r \in [0,1]^I$ by $\bar{v}^r_i = v_r(i)$ for all $i \in I$. That is, $\bar{v}^r$ is a representation of the volume of the items in the r-th dimension as a vector. For every $\bar{t} \in \mathcal{T}$ define

$$\mathcal{S}_{\text{large},\bar{t}} = \left\{ \mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}} \;\middle|\; \bar{a} \in \mathcal{G} \right\} \qquad \text{and} \qquad \mathcal{S}_{\text{small},\bar{t}} = \left\{ \mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r \;\middle|\; r = 1, 2, \ldots, d \right\}.$$

Finally, define $\mathcal{S} = \bigcup_{\bar{t} \in \mathcal{T}} \left(\mathcal{S}_{\text{large},\bar{t}} \cup \mathcal{S}_{\text{small},\bar{t}} \right)$.

Claim 3.14. *It holds $|\mathcal{S}| \leq \kappa(\delta)$.*

Proof. By a simple counting argument,

$$\begin{aligned}
|\mathcal{S}| &\leq \sum_{\bar{t} \in \mathcal{T}} \left(\left|\mathcal{S}_{\text{large},\bar{t}}\right| + \left|\mathcal{S}_{\text{small},\bar{t}}\right| \right) \\
&\leq |\mathcal{T}| \cdot (|\mathcal{G}| + d) \\
&\leq \exp\left(\delta^{-5\delta^{-1}}\right) \left((d \cdot h)^d + d \right) \\
&\leq \exp\left(\delta^{-6 \cdot \delta^{-1}}\right) \\
&\leq \exp\left(\exp\left(-5 \cdot \delta^{-1} \cdot \ln(\delta)\right)\right) \\
&\leq \kappa(\delta) \ .
\end{aligned}$$

The third inequality uses (30) and the forth inequality uses $d \leq \delta^{-1}$. ◇

We are left to show the constructed structure $\mathcal{S}$ satisfies the condition in Lemma 3.5. The following claims provide some basic properties which will assist us in achieving this goal.

Claim 3.15. *Let $\bar{t} \in \mathcal{T}$ and let $C \in \mathcal{C}$ be such that $C \subseteq L \cap R$ and $\mathsf{T}(C) \leq \bar{t}$. That is, for all $\bar{a} \in \mathcal{G}$ it holds that $\mathsf{T}_{\bar{a}}(C) \leq \bar{t}_{\bar{a}}$. Then $v(C) \leq \tilde{v}(\bar{t})$.*

Proof. For $r = 1, \ldots, d$ it holds that

$$\begin{aligned}
v_r(C) &= \sum_{\bar{a} \in \mathcal{G}} \sum_{i \in C \cap I_{\bar{a}}} v_r(i) \\
&\leq \sum_{\bar{a} \in \mathcal{G}} \sum_{i \in C \cap I_{\bar{a}}} \tilde{v}_r(\bar{a}) \\
&= \sum_{\bar{a} \in \mathcal{G}} \mathsf{T}_{\bar{a}}(C) \cdot \tilde{v}_r(\bar{a}) \\
&\leq \sum_{\bar{a} \in \mathcal{G}} \bar{t}_{\bar{a}} \cdot \tilde{v}_r(\bar{a}) \ = \bar{v}_r(\bar{t}) \ .
\end{aligned}$$

The first inequality holds since $v_r(i) \leq \tilde{v}_r(\bar{a})$ for every $i \in I_{\bar{a}}$ by (28) and (29). The second inequality follows from the assumptions of the claim. The last equality follows from the definition of $\tilde{v}(\bar{t})$ in (31). ◇

Claim 3.16. *Let $\bar{t} \in \mathcal{T}$ and $C \in \mathcal{C}$ such that $C \subseteq R \cap L$ and $\mathsf{T}(C) = \bar{t}$. Then $v_r(C) \geq \tilde{v}_r(C) - \frac{\delta}{2}$ for $r = 1, \ldots, d$.*

Proof. For $r = 1, \ldots, d$ it holds that

$$\begin{aligned} v_r(C) &= \sum_{\bar{a} \in \mathcal{G}} \sum_{\bar{a} \in C \cap I_{\bar{a}}} v_r(i) \\ &\geq \sum_{\bar{a} \in \mathcal{G}} \sum_{\bar{a} \in C \cap I_{\bar{a}}} \left(\tilde{v}_r(\bar{a}) - \frac{\delta^2}{2 \cdot d} \right) \\ &= \tilde{v}_r(\bar{t}) - |C \cap L| \cdot \frac{\delta^2}{2d} \\ &\geq \tilde{v}_r(\bar{t}) - \frac{\delta}{2} \ . \end{aligned}$$

The first inequality follows from (28) and (29). The last inequality holds, as $|C \cap \text{Ł}| \leq d \cdot \delta^{-1}$. $\diamond$

Claim 3.17. *Let $\ell \in \{1, \ldots, s\}$ and $\bar{t} = \mathsf{T}(B_\ell)$. Then $v_r(B_\ell \setminus L) \leq 1 - \tilde{v}_r(\bar{t}) - \frac{\delta}{2}$ for $r = 1, \ldots, d$.*

Proof. For $r = 1, \ldots, d$ we have

$$v_r(B_\ell \setminus L) = v_r(B_\ell) - v_r(B_\ell \cap L) \leq 1 - \delta - \left(\tilde{v}_r(\bar{t}) - \frac{\delta}{2} \right) = 1 - \tilde{v}_r(\bar{t}) - \frac{\delta}{2} \ .$$

The inequality holds since B_ℓ has δ-full slack and by Claim 3.16. $\diamond$

The following is an immediate consequence of Claim 3.17.

Corollary 3.18. *For all $\bar{t} \in \mathcal{T}$ and $r \in \{1, 2, \ldots, d\}$ it holds that $\tilde{v}_r(\bar{t}) \leq 1 - \frac{\delta}{2}$.*

Let $Q \subseteq R$ and $\gamma \in (0, 1)$ be such that

$$\forall \bar{u} \in \mathcal{S} : \qquad \mathbb{1}_Q \cdot \bar{u} \leq \gamma \cdot \mathbb{1}_R \cdot \bar{u} + \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT}(I, v) \cdot \mathsf{tol}(\bar{u}) \ . \tag{32}$$

To complete the proof, we need to show that $\mathrm{OPT}(Q, v) \leq \gamma(1 + d \cdot \delta) \cdot s + \delta^{10} \cdot \mathrm{OPT} + \kappa(\delta)$. Towards this end, we will construct a separate packing of $Q \cap (L_{\bar{t}} \cup S_{\bar{t}})$ for every $\bar{t} \in \mathcal{T}$.

Define the *prevalence* of type $\bar{t} \in \mathcal{T}$ by $p_{\bar{t}} = |\{\ell \in \{1, \ldots, s\} \mid \mathsf{T}(B_\ell) = \bar{t}\}|$. That is, $p_{\bar{t}}$ is the number of configuration among $B_1, \ldots, B_s$ of type $\bar{t}$. For every $\bar{t} \in \mathcal{T}$, define

$$\eta_{\bar{t}} = \left\lceil \gamma \cdot p_{\bar{t}} + \frac{\delta^{15}}{\kappa(\delta)} \cdot \mathrm{OPT} \right\rceil \ . \tag{33}$$

We will show that $\mathrm{OPT}(Q \cap (L_{\bar{t}} \cup S_{\bar{t}}) \setminus X_{\bar{t}}, v) \leq \eta_{\bar{t}}$ where $X_{\bar{t}}$ is a set that satisfies $\mathrm{OPT}(X_t, v) \leq \delta \cdot \eta_{\bar{t}}$.

Claim 3.19. *For every $\bar{t} \in \mathcal{T}$ there exists $D_1^{\bar{t}}, \ldots, D_{\eta_{\bar{t}}}^{\bar{t}} \subseteq I$ such that $\bigcup_{\ell=1}^{\eta_{\bar{t}}} D_\ell^{\bar{t}} = Q \cap L_{\bar{t}}$ and $v(D_\ell^{\bar{t}}) \leq \tilde{v}(\bar{t})$ for $\ell = 1, \ldots, \eta_{\bar{t}}$.*

By Claim 3.19 we can pack the items in $Q \cap L_{\bar{t}}$ into $\eta_{\bar{t}}$ configuration with volume at most $\tilde{v}(\bar{t})$. The unused volume of $1 - \tilde{v}_r(\bar{t})$ in each coordinate $r = 1, \ldots, d$ will be used to pack the set small items $Q \cap S_{\bar{t}}$.

Proof of Claim 3.19. Let $\mathcal{G}_{\bar{t}} = \{\bar{a} \in \mathcal{G} \mid \bar{t}_{\bar{a}} \neq 0\}$. For every $\bar{a} \in \mathcal{G} \setminus \mathcal{G}_{\bar{t}}$, we have $\bar{t}_{\bar{a}} = 0$, and therefore $L_{\bar{t}} \cap I_{\bar{a}} = \emptyset$ (configurations of type $\bar{t}$ do not contain items from $I_{\bar{a}}$, and $L_{\bar{t}}$ is a set of items in configurations of type $\bar{t}$). Thus $Q \cap L_{\bar{t}} \cap I_{\bar{a}} = \emptyset$ for all $\bar{a} \in \mathcal{G} \setminus \mathcal{G}_{\bar{t}}$.

For all $\bar{a} \in \mathcal{G}_{\bar{t}}$ it holds that $\mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}} \in \mathcal{S}_{\text{large},\bar{t}} \subseteq \mathcal{S}$. Thus, by (32) we have

$$|Q \cap I_{\bar{a}} \cap L_{\bar{t}}| \;=\; \mathbb{1}_Q \cdot \mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}} \;\leq\; \gamma \cdot \mathbb{1}_R \cdot \mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}} + \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT}(I, v) \cdot \mathsf{tol}\left(\mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}}\right) \;. \tag{34}$$

Furthermore, for all $C \in \mathcal{C}$ it holds that $\sum_{i \in C} \left(\mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}}\right)_i \leq \sum_{i \in C} \mathbb{1}_{i \in L} \leq d \cdot \delta^{-1} \leq \delta^{-2}$, thus $\mathsf{tol}(\mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}}) \leq \delta^{-2}$. By plugging the last inequality into (34) we obtain,

$$|Q \cap I_{\bar{a}} \cap L_{\bar{t}}| \;\leq\; \gamma \cdot \mathbb{1}_R \cdot \mathbb{1}_{I_{\bar{a}} \cap L_{\bar{t}}} + \frac{\delta^{18}}{\kappa(\delta)} \cdot \mathrm{OPT}(I, v) \;\leq\; \gamma \cdot |R \cap I_{\bar{a}} \cap L_{\bar{t}}| + \frac{\delta^{18}}{\kappa(\delta)} \cdot \mathrm{OPT}(I, v) \;. \tag{35}$$

Observe that

$$|R \cap I_{\bar{a}} \cap L_{\bar{t}}| \;=\; \sum_{\ell \in [s] \text{ s.t. } \mathsf{T}(B_\ell) = \bar{t}} |B_\ell \cap I_{\bar{a}}| = \sum_{\ell \in [s] \text{ s.t. } \mathsf{T}(B_\ell) = \bar{t}} \bar{t}_{\bar{a}} = p_{\bar{t}} \cdot \bar{t}_{\bar{a}} \;. \tag{36}$$

By (35) and (36), it holds that

$$|Q \cap I_{\bar{a}} \cap L_{\bar{t}}| \;\leq\; \gamma \cdot p_{\bar{t}} \cdot \bar{t}_{\bar{a}} + \frac{\delta^{18}}{\kappa(\delta)} \cdot \mathrm{OPT}(I, v) \leq \bar{t}_{\bar{a}} \cdot \eta_{\bar{t}}.$$

Therefore, for every $\bar{a} \in \mathcal{G}_{\bar{t}}$ we can partition $Q \cap I_{\bar{a}} \cap L_{\bar{t}}$ into $\eta_{\bar{t}}$ sets $D^{\bar{t}}_{1,\bar{a}}, \ldots, D^{\bar{t}}_{\eta_{\bar{t}},\bar{a}}$ such that $\left|D^{\bar{y}}_{\ell,\bar{a}}\right| \leq \bar{t}_{\bar{a}}$ (we allow sets in the partition to be empty). Define sets $D^{\bar{t}}_1, \ldots, D^{\bar{t}}_{\eta_{\bar{t}}}$ by $D^{\bar{t}}_\ell = \bigcup_{\bar{a} \in \mathcal{G}_{\bar{t}}} D^{\bar{t}}_{\ell,\bar{a}}$ for all $\ell = 1, 2, \ldots, \eta_{\bar{t}}$. It follows that

$$\bigcup_{\ell=1}^{\eta_{\bar{t}}} D^{\bar{t}}_\ell \;=\; \bigcup_{\ell=1}^{\eta_{\bar{t}}} \bigcup_{\bar{a} \in \mathcal{G}_{\bar{t}}} D^{\bar{t}}_{\ell,\bar{a}} \;=\; \bigcup_{\bar{a} \in \mathcal{G}_{\bar{t}}} (Q \cap I_{\bar{a}} \cap L_{\bar{t}}) = Q \cap L_{\bar{t}} \;.$$

For all $\bar{a} \in \mathcal{G} \setminus \mathcal{G}_{\bar{t}}$ and $\ell = 1, \ldots, \eta_{\bar{t}}$ it holds that $\mathsf{T}_{\bar{a}}(D^{\bar{t}}_\ell) = \left|D^{\bar{t}}_\ell \cap I_{\bar{a}}\right| = 0 = \bar{t}_{\bar{a}}$. Furthermore, for all $\bar{a} \in \mathcal{G}_{\bar{t}}$ and $\ell = 1, \ldots, \eta_{\bar{t}}$ it holds that t $\mathsf{T}_{\bar{a}}(D^{\bar{t}}_\ell) = \left|D^{\bar{t}}_\ell \cap I_{\bar{a}}\right| = \left|D^{\bar{t}}_{\ell,\bar{a}}\right| \leq \bar{t}_{\bar{a}}$. Thus, $\mathsf{T}(D^{\bar{t}}_\ell) \leq \bar{t}$ for all $\ell = 1, \ldots, \eta_{\bar{t}}$. By Claim 3.15, it follows that $v(D^{\bar{t}}_\ell) \leq \tilde{v}(\bar{t})$ for all $\ell = 1, \ldots, \eta_{\bar{t}}$. ◇

While Claim 3.19 handles the large items in Q, the next claim deals with the small items in Q.

Claim 3.20. *For all $\bar{t} \in \mathcal{T}$ there exists $F^{\bar{t}}_1, \ldots, F^{\bar{t}}_{\eta_{\bar{t}}} \subseteq I$ and $X_{\bar{t}} \subseteq I$ such that*

- $\bigcup_{\ell=1}^{\eta_{\bar{t}}} F^{\bar{t}}_\ell = (Q \cap S_{\bar{t}}) \setminus X_{\bar{t}}$,
- $\mathrm{OPT}(X_{\bar{t}}, v) \leq \delta \cdot d \cdot \eta_{\bar{t}} + 1$,
- *and* $v_r(F^{\bar{t}}_\ell) \leq 1 - \tilde{v}_r(\bar{t})$ *for all* $\ell = 1, \ldots, \eta_{\bar{t}}$ *and* $r = 1, \ldots, d$.

Proof. For all $r = 1, \ldots, d$ it holds that $\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r \in \mathcal{S}$. Thus, by (32) it holds that

$$v_r(Q \cap S_{\bar{t}}) = \mathbb{1}_Q \cdot \left(\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r\right) \;\leq\; \gamma \cdot \mathbb{1}_R \cdot \left(\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r\right) + \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r) \;. \tag{37}$$

For all $C \in \mathcal{C}$ it holds that $\sum_{i\in C} \left(\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r\right)_i \leq \sum_{i\in C} v_r(i) \leq 1$, hence $\mathsf{tol}(\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r) \leq 1$. Thus, we can rewrite (37) as

$$v_r(Q \cap S_{\bar{t}}) \;\leq\; \gamma \cdot \mathbb{1}_R \cdot \left(\mathbb{1}_{S_{\bar{t}}} \wedge \bar{v}^r\right) + \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT} \;=\; \gamma \cdot v_r(R \cap S_{\bar{t}}) + \frac{\delta^{20}}{\kappa(\delta)} \cdot \mathrm{OPT} \;. \tag{38}$$

By the definition of $S_{\bar{t}}$ we also have

$$v_r(R \cap S_{\bar{t}}) = \sum_{\ell \in [s] \text{ s.t.} \mathsf{T}(B_\ell) = \bar{t}} v_r(B_\ell \setminus L) \leq \sum_{\ell \in [s] \text{ s.t.} \mathsf{T}(B_\ell) = \bar{t}} \left(1 - \tilde{v}_r(\bar{t}) - \frac{\delta}{2}\right) \leq p_{\bar{t}} \cdot \left(1 - \tilde{v}_r(\bar{t})\right), \tag{39}$$

where the first inequality follows from Claim 3.17. By (38) and (39) we have

$$v_r(Q \cap S_{\bar{t}}) \;\leq\; \gamma p_{\bar{t}} \left(1 - \tilde{v}_r(\bar{t})\right) + \frac{\delta^{20}}{\kappa(\delta)} \mathrm{OPT} \;\leq\; \left(1 - \tilde{v}_r(\bar{t})\right) \cdot \left(\gamma p_{\bar{t}} + \frac{\delta^{15}}{\kappa(\delta)} \mathrm{OPT}\right) \leq \left(1 - \tilde{v}_r(\bar{t})\right) \cdot \eta_{\bar{t}}, \tag{40}$$

where the second inequality follows from Corollary 3.18.

Our construction utilizes integrality properties of the polytope P defined by

$$P = \left\{ \bar{\mu} \in [0,1]^{Q \cap S_{\bar{t}} \times [\eta_{\bar{t}}]} \;\middle|\; \begin{array}{ll} \sum_{\ell=1}^{\eta_{\bar{t}}} \bar{\mu}_{i,\ell} = 1 & \forall i \in Q \cap S_{\bar{t}} \\ \sum_{i \in Q \cap S_{\bar{t}}} v_r(i) \cdot \bar{\mu}_{i,\ell} \leq 1 - \tilde{v}_r(\bar{t}) & \forall r \in [d],\; \ell \in [\eta_{\bar{t}}], \end{array} \right\} . \tag{41}$$

That is, an entry in P is a vector with entries of the form $\bar{\mu}_{i,\ell}$, where $i \in Q \cap S_{\bar{t}}$ and $\ell \in \{1, \ldots, \eta_{\bar{t}}\}$. The entry $\bar{\mu}_{i,\ell}$ can be interpreted as the fractional assignment of the item i to the ℓ-th bin. The first constraint in (41) ensures all the items are fully assigned, and the second constraint enforces an upper bound on the total volume of items assigned to a specific bin in each coordinate. It is well-known (see, e.g., [BEK16]) that a vertex of P contains at most $d \cdot \eta_{\bar{t}}$ fractional entries. Formally, if $\bar{\mu}^* \in P$ is a vertex of P then $\left|\left\{(i,\ell) \in Q \cap S_t \times \{1, \ldots, \eta_{\bar{t}}\} \mid \bar{\mu}^*_{i,\ell} \in (0,1)\right\}\right| \leq d \cdot \eta_{\bar{t}}$.

In order to exploit the above-mentioned property of P, we first need to show $P \neq \emptyset$. Define $\bar{x} \in [0,1]^{Q \cap S_{\bar{t}} \times \{1, \ldots, \eta_{\bar{t}}\}}$ by $\bar{x}_{i,\ell} = \frac{1}{\eta_{\bar{t}}}$ for all $i \in Q \cap S_{\bar{t}}$ and $\ell = 1, \ldots, \eta_{\bar{t}}$. For all $i \in Q \cap S_{\bar{t}}$ it holds that

$$\sum_{\ell=1}^{\eta_{\bar{t}}} \bar{x}_{i,\ell} = \sum_{\ell=1}^{\eta_{\bar{t}}} \frac{1}{\eta_{\bar{t}}} = 1 \;. \tag{42}$$

Furthermore, for every $\ell = 1, \ldots, \eta_{\bar{t}}$ and $r = 1, \ldots, d$ we have

$$\sum_{i \in Q \cap S_{\bar{t}}} v_r(i) \cdot \bar{x}_{i,\ell} = \sum_{i \in Q \cap S_{\bar{t}}} v_r(i) \cdot \frac{1}{\eta_{\bar{t}}} = \frac{1}{\eta_{\bar{t}}} \cdot v_r(Q \cap S_{\bar{t}}) \leq 1 - \tilde{v}_r(\bar{t}), \tag{43}$$

where the last inequality follows from (40). By (42) and (43) we have $\bar{x} \in P$, and thus $P \neq \emptyset$.

Therefore, there exists a vertex $\bar{\mu}^*$ of the polytope P and it holds that

$$\left|\left\{(i,\ell) \in Q \cap S_t \times \{1, \ldots, \eta_{\bar{t}}\} \mid \bar{\mu}^*_{i,\ell} \in (0,1)\right\}\right| \leq d \cdot \eta_{\bar{t}} \;.$$

Define $X_{\bar{t}} = \left\{ i \in Q \cap S_{\bar{t}} \;\middle|\; \exists \ell \in \{1, \ldots, s\} : \; \bar{\mu}^*_{i,\ell} \in (0,1) \right\}$. It thus holds that $|X_{\bar{t}}| \leq d \cdot \eta_{\bar{t}}$. Since all items in $X_{\bar{t}}$ are small, it holds that every subset of δ^{-1} items of $X_{\bar{t}}$ form a configuration, thus $\mathrm{OPT}(X_{\bar{t}}, v) \leq \delta |X_{\bar{t}}| + 1 \leq \delta \cdot d \cdot \eta_{\bar{t}} + 1$.

For $\ell = 1, \ldots, \eta_{\bar{t}}$ define $F_\ell^{\bar{t}} = \left\{ i \in Q \cap S_{\bar{t}} \;\middle|\; \bar{\mu}^*_{i,\ell} = 1 \right\}$. As $\bar{\mu}^* \in P$ (41) it holds follows that $v_r(F_\ell^{\bar{t}}) \leq \sum_{i \in Q \cap S_{\bar{t}}} v_r(i) \cdot \bar{\mu}^*_{i,\ell} \leq 1 - \tilde{v}_r(\bar{t})$ for all $r = 1, \ldots, d$. Furthermore,

$$\bigcup_{\ell=1}^{\eta_{\bar{t}}} F_\ell^{\bar{t}} = \{ i \in Q \cap S_{\bar{t}} \mid \forall \ell = 1, \ldots, \eta_{\bar{t}} : \; \bar{\mu}^*_{i,\ell} \in \{0,1\} \} = (Q \cap S_{\bar{t}}) \setminus X_{\bar{t}},$$

which completes the proof of the claim. $\diamond$

For every $\bar{t}$ let $D_1^{\bar{t}}, \ldots, D^{\bar{t}}{}_{,\eta_{\bar{t}}}$ be the sets from Claim 3.19 and let $X_{\bar{t}}$ and $F_1^{\bar{t}}, \ldots, F_{\eta_{\bar{t}}}^{\bar{t}}$ be the sets from Claim 3.20. It follows that $v_r(D_\ell^{\bar{t}} \cup F_\ell^{\bar{t}}) \leq \tilde{v}_r(\bar{t}) + 1 - \tilde{v}_r(\bar{t}) = 1$ for all $\ell = 1, \ldots, \eta_{\bar{t}}$ and $r = 1, \ldots, d$. Thus $D_\ell^{\bar{t}} \cup F_\ell^{\bar{t}} \in \mathcal{C}$ for all $\ell = 1, \ldots, \eta_{\bar{t}}$. It also holds that $\bigcup_{\ell=1}^{\eta_{\bar{t}}} \left(D_\ell^{\bar{t}} \cup F_\ell^{\bar{t}} \right) = (Q \cap (L_{\bar{t}} \cup S_{\bar{t}})) \setminus X_{\bar{t}}$. Therefore,

$$\mathrm{OPT}\left(Q \cap (L_{\bar{t}} \cup S_{\bar{t}}), v\right) \;\leq\; \mathrm{OPT}\left((Q \cap (L_{\bar{t}} \cup S_{\bar{t}})) \setminus X_{\bar{t}}, v\right) + \mathrm{OPT}(X_{\bar{t}}, v) \;\leq\; \eta_{\bar{t}} + \delta \cdot d \cdot \eta_{\bar{t}} + 1,$$

and thus,

$$\begin{aligned}
\mathrm{OPT}(Q, v) \leq\; & \sum_{\bar{t} \in \mathcal{T}} \mathrm{OPT}\left(Q \cap (L_{\bar{t}} \cup S_{\bar{t}}), v\right) \\
\leq\; & \sum_{\bar{t} \in \mathcal{T}} (\eta_{\bar{t}} + \delta \cdot d \cdot \eta_{\bar{t}} + 1) \\
\leq\; & |\mathcal{T}| + (1 + \delta \cdot d) \sum_{\bar{t} \in \mathcal{T}} \eta_{\bar{t}} \\
=\; & |\mathcal{T}| + (1 + \delta \cdot d) \sum_{\bar{t} \in \mathcal{T}} \left\lceil \gamma \cdot p_{\bar{t}} + \frac{\delta^{15}}{\kappa(\delta)} \cdot \mathrm{OPT} \right\rceil \\
\leq\; & 3 \cdot |\mathcal{T}| + (1 + \delta \cdot d) \sum_{\bar{t} \in \mathcal{T}} \left(\gamma \cdot p_{\bar{t}} + \frac{\delta^{15}}{\kappa(\delta)} \cdot \mathrm{OPT} \right) \\
=\; & 3 \cdot |\mathcal{T}| + \gamma \cdot (1 + \delta \cdot d) \sum_{\bar{t} \in \mathcal{T}} p_{\bar{t}} + (1 + \delta \cdot d) \cdot |\mathcal{T}| \cdot \frac{\delta^{15}}{\kappa(\delta)} \cdot \mathrm{OPT} \\
\leq\; & \kappa(\delta) + \gamma \cdot (1 + \delta \cdot d) \cdot s + \delta^{10} \cdot \mathrm{OPT} \;.
\end{aligned}$$

The first equality follows from (33). The last inequality uses (30) and $d \leq \delta^{-1}$. $\square$

4 Asymptotic $\left(\frac{4}{3} + \varepsilon\right)$ approximation for 2VBP

In this section we prove Lemma 1.9. That is, we show that Algorithm 2 is a randomized asymptotic $\left(\frac{4}{3} + \varepsilon\right)$-approximation algorithm for 2VBP. The analysis of the algorithm utilizes a variant of the Configuration-LP (1) in which each item $i \in I$ has a *demand* $\bar{d}_i \in [0,1]$. That is, given a 2VBP instance and for every *demand* vector $\bar{d} \in [0,1]^I$ define

$$\begin{aligned}
\text{Demand-LP}(\bar{d}): \quad & \min & & \sum_{C \in \mathcal{C}} \bar{x}_C, \\
& \forall i \in I: & & \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = \bar{d}_i, \\
& \forall C \in \mathcal{C}: & & \bar{x}_C \geq 0 \;.
\end{aligned} \tag{44}$$

Observe that for every $S \subseteq I$ it holds that $\text{LP}(S)$ is identical to $\text{Demand-LP}(\mathbb{1}_S)$. We use $\text{OPT}_f(\bar{d})$ to denote the value of an optimal solution for $\text{Demand-LP}(\bar{d})$

We extend the definition of configuration to allow multiple occurrences of items. Let (I, v) be a 2VBP instance. A multi-set over I is a function $C : I \to \mathbb{N}$. For $i \in I$ we say that $i \in C$ if $C(i) > 0$. A *multi-configuration* is a multi-set C over I such that $v(C) = \sum_{i \in I} C(i) \cdot v(i) \leq (1, 1)$. We use $\mathcal{C}^*$ to denote the set of all multi-configurations. We identify the set $C \subseteq I$ with the multi-set C' in which $C'(i) = C(i)$.

Given $\bar{x} \in [0,1]^{\mathcal{C}}$ ($\bar{x} \in [0,1]^{\mathcal{C}^*}$) the *coverage* of $\bar{x}$ is the vector $\bar{y} \in [0,1]^I$ defined by $\bar{y}_i = \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i)$ ($\bar{y}_i = \sum_{C \in \mathcal{C}^*} \bar{x}_C \cdot C(i)$) for every $i \in I$. We say that $\bar{y} \in [0,1]^I$ is *small-items integral* if $\bar{y}_i \in \{0, 1\}$ for any $i \in I \setminus L$. Similarly, we say that $\bar{x} \in [0,1]^{\mathcal{C}}$ ($\bar{x} \in [0,1]^{\mathcal{C}^*}$) is *small-items integral* if its coverage is small-items integral.

Recall that $\text{OPT}(I, v)$ is the minimum solution size for the instance (I, v). Our analysis relies on the existence of "linear structures".

Definition 4.1 (Linear Structure)**.** *Let $\delta, K > 0$. Let (I, v) be a δ-2VBP instance, let $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$, and let $\bar{w} \in [0,1]^I$ be the coverage of $\bar{\lambda}$. A (δ, K)-linear structure of $\bar{\lambda}$ is a subset $\mathcal{S} \subseteq \mathbb{R}^I_{\geq 0}$ of size at most K which satisfies the following property. For any small-items integral vector $\bar{z} \in [0,1]^I$ and $\beta \in \left[\delta^5, 1\right]$ such that* $\text{supp}(\bar{z}) \subseteq \text{supp}(\bar{w})$ *and*

$$\bar{z} \cdot \bar{u} \leq \beta \cdot \bar{w} \cdot \bar{u} + \frac{1}{K^{10}} \cdot \text{OPT}(I, v) \cdot \mathsf{tol}(\bar{u}), \tag{45}$$

for all $\bar{u} \in \mathcal{S}$, it holds that $\text{OPT}_f(\bar{z}) \leq \beta \cdot (1 + 10\delta) \cdot \|\bar{\lambda}\| + K + \delta^{10} \cdot \text{OPT}(I, v)$.

Observe that a linear structure has properties similar to a weak structure (Lemma 3.5). Intuitively, a linear structure implies that if a demand vector $\bar{z}$ satisfies a 'small' number of constraints with respect to β (where K is a constant, as defined in Lemma 4.2) then we obtain a decrease in $\text{OPT}_f(\bar{z})$ by factor of β. While linear structures do not necessarily exist for arbitrary vectors $\bar{\lambda}$, we show that such structures exist for vectors which only select configurations with *slack*. We say that $C \in \mathcal{C}^*$ *has δ-slack in dimension $d \in \{1, 2\}$* if $v_d(C) \leq 1 - \delta$. We say that $C \in \mathcal{C}^*$ *has δ-slack* if there is $d \in \{1, 2\}$ such that C has δ-slack in dimension d. Finally, we say that $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ is *with δ-slack* if every configuration $C \in \text{supp}(\bar{\lambda})$ has δ-slack.

Lemma 4.2 (Structural Property)**.** *Let (I, v) be a δ-2VBP instance, where $\delta \in (0, 0.1)$, and $\delta^{-1} \in \mathbb{N}$. There is a set $\mathcal{S}^* \subseteq \mathbb{R}^I_{\geq 0}$ such that $|\mathcal{S}^*| \leq \varphi(\delta) \cdot |L|^4$, where $\varphi(\delta) = \exp\left(\delta^{-20}\right)$, which satisfies the following property. For any small-items integral $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ with δ-slack, there is a $(\delta, \varphi(\delta))$-linear structure $\mathcal{S}$ of $\bar{\lambda}$ where for all $\bar{u} \in \mathcal{S}$: if* $\text{supp}(\bar{u}) \cap L \neq \emptyset$ *then $\bar{u} \in \mathcal{S}^*$.*

The proof of the lemma (given in Section 4.2) uses some of the structural features shown by Bansal et al. [BEK16], along with the recent concept of fractional grouping, adopted from Fairstein et al. [FKS21]. While the set $\mathcal{S}^*$ does not limit the number of structures which may be generated by the lemma, it limits the set of vectors these structures may use. This attribute is crucial for our analysis (specifically, in the proof of Lemma 4.16).

To show the existence of linear structure we often need to convert an arbitrary configuration to a vector $\bar{\lambda}$ with a slack. To this end, we use the following definition and lemmas.

Definition 4.3. *Given $C \in \mathcal{C}$ and $\psi \geq 1$, we say that $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ is a ψ-*relaxation *of C if the following conditions simultaneously hold:*

1. *$\bar{\lambda}$ is with δ-slack,*

2. $\|\bar{\lambda}\| \leq \psi$,

3. and $\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = C(i)$ for every $i \in I$.

Lemma 4.4. *Let $\delta \in (0, 0.1)$ be such that $\delta^{-1} \in \mathbb{N}$ and let (I, v) be a δ-2VBP instance. Then for any $C \in \mathcal{C}_0$, there is a $(1 + 4\delta)$-relaxation of C.*

Lemma 4.5. *Let $\delta \in (0, 0.1)$ and let (I, v) be a δ-2VBP instance. Then for any $h = 2, \ldots, 2\delta^{-1}$ and $C \in \mathcal{C}_h$ there is an $\frac{h}{h-1}$-relaxation of C.*

Lemma 4.6. *Let $\delta \in (0, 0.1)$, let (I, v) be a δ-2VBP instance, and let $C \in \mathcal{C}$ such that $v(C) \leq (\delta, \delta)$. Then there is a 4δ-relaxation of C.*

The proofs of Lemma 4.4, Lemma 4.5, and Lemma 4.6 are given in Section 4.3. Some of the statements and techniques used in the proofs can be viewed as variants of [BEK16, Lemma 5.3]. We proceed to the analysis of Algorithm 2 in Section 4.1. The PTAS for the Matching Configuration LP (4) (Lemma 1.8) is given in Section 4.4.

4.1 The Analysis of Match&Round

Throughout this section, we fix a δ-2VBP instance (I, v) and $\delta \in (0, 0.1)$ such that $\delta^{-1} \in \mathbb{N}$. Thus, notations such as ρ_j, S_j C^j_ℓ, and $\mathcal{M}$ refer to the corresponding variables in the execution of Algorithm 2 (and the call to Algorithm 1 as part of its execution), with (I, v) as its input and δ as the parameter. We also use $\varphi(\delta) = \exp(\delta^{-20})$ as in Lemma 4.2 and $\mathrm{OPT} = \mathrm{OPT}(I, v)$. We commonly use $k = \lceil \ln_{1-\delta}(\delta) \rceil \leq \delta^{-2}$.

The core of the analysis is in Section 4.1.1, in which we derive a bound on the number of configurations sampled by Algorithm 1. Section 4.1.2 gives the proof of Lemma 1.9. The analysis involves the use of several concentration bounds whose proofs are simple yet technical. To avoid diversion from the main flow of the analysis, we defer the proofs of the concentration bounds to Section 4.1.3.

We use the probabilistic space $(\Omega, \mathcal{F}, \Pr)$ as defined Section 2. Recall that Lemma 2.6 provides an upper bound on ρ^*, the size of the solution returned by First-Fit in Line 6 of Algorithm 1. Also, observe that $\mathbb{E}[|\mathcal{M}|] = (1 - \delta^4) \cdot \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C}_2}$ (recall $\mathcal{C}_2$ is defined in (2)). We use the concentration bounds of Chekuri, Vondrák and Zenklusen [CVZ11] to show that, with high probability, $|\mathcal{M}|$ is close to its expectation.

Lemma 4.7. *It holds that $|\mathcal{M}| \leq \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C}_2} + \delta^2 \cdot \mathrm{OPT}$ with probability at least $1 - \exp\left(-\delta^{10} \cdot \mathrm{OPT}\right)$.*

The proof of the lemma is given in Section 4.1.3.

The size of the solution returned by Algorithm 2 is $|\mathcal{M}| + \sum_{j=1}^{k} \rho_j + \rho^*$. As Lemma 2.6 and Lemma 4.7 give upper bounds for $|\mathcal{M}|$ and ρ^*, it remains to derive an upper bound on $\sum_{j=1}^{k} \rho_j$, the total number of configurations sampled by Iterative Randomized Rounding.

4.1.1 A Refined Analysis of the Iterattve Rounding

Our analysis relies on the key notion of "untouched" configurations. Recall the sets of configurations $\mathcal{C}_j$ were define in (2), and $\mathcal{C}_0 = \mathcal{C} \setminus \left(\bigcup_{h=2}^{2 \cdot \delta^{-1}} \mathcal{C}_h \right)$. For iteration $j \in \{0, 1, \ldots, k\}$, define the set of *untouched configurations* as

$$U_j = \{C \in \mathcal{C} \mid C \cap S_j \notin \mathcal{C}_0\} = \{C \in \mathcal{C} \mid v(C \cap S_j \cap L) > (1 - \delta, 1 - \delta)\} \quad .$$

Since $S_0 \supseteq S_1 \supseteq \ldots \supseteq S_k$, it follows that $U_0 \supseteq U_1 \supseteq \ldots \supseteq U_k$. We denote by $T_0 = \mathcal{C} \setminus U_0$ the initial set of *touched* configurations, and by $T_j = U_{j-1} \setminus U_j$ the configurations that become touched in iteration j, for $j = 1, \ldots, k$. Observe that $\mathcal{C}_0 \subseteq T_0$. We refine the sets U_j and T_j by defining $U_{j,h} = U_j \cap \mathcal{C}_h$ and $T_{j,h} = T_j \cap \mathcal{C}_h$ for $j = 0, \ldots, k$ and $h = 0, \ldots, 2 \cdot \delta^{-1}$.

Intuitively, we view configurations in $\mathcal{C}_0$ as "easy" compared to configurations in $\mathcal{C} \setminus \mathcal{C}_0$. Indeed, we can construct linear structures only for configurations with a slack (Lemma 4.2), and a slack can be obtained with negligible overhead for configurations in $\mathcal{C}_0$. Thus, configurations in U_j "remain difficult" after iteration j, while configurations in T_j "become easy" in iteration j. Observe that

$$\sum_{j=1}^{k} \rho_j \leq k + \alpha(1+\delta^2) \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j}) \leq k + (1+2\delta)\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j}), \tag{46}$$

where the first inequality uses $\rho_j = \lceil \alpha z_j \rceil \leq \alpha(1+\delta^2)\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}}) + 1$, and the second inequality uses $\alpha(1+\delta^2) \leq (1+2\delta)\delta$. Next, we derive an upper bound on $\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j})$. By (46), this would imply a bound on $\sum_{j=1}^{k} \rho_j$, the number of configurations sampled by Algorithm 1.

Recall that $\bar{x}^0$ is the solution for MLP found in Line 2 of Algorithm 2. We define $\bar{x}^* \in [0,1]^{\mathcal{C}}$ by

$$\bar{x}^*_C = \sum_{C' \in U_0 \setminus \mathcal{C}_2 \text{ s.t. } C' \cap L = C} \bar{x}^0_{C'}$$

for each $C \in \mathcal{C}$. Inzuitively, $\bar{x}^*$ can be viewed as selecting all the configurations in $U_0 \setminus \mathcal{C}_2$ as in $\bar{x}^0$, and then discarding the small items. Since U_0 is $\mathcal{F}_0$-measurable and $\bar{x}^0$ is $\mathcal{F}_{-1}$-measurable, it follows that $\bar{x}^*$ is $\mathcal{F}_0$-measurable. It can be easily verified that $\bar{x}^* \cdot \mathbb{1}_{\mathcal{C}_h} = \bar{x}^0 \cdot \mathbb{1}_{U_{0,h}}$ for every $3 \leq h \leq 2 \cdot \delta^{-1}$ and $\bar{x}^* \cdot \mathbb{1}_{\mathcal{C}_0} = \bar{x}^* \cdot \mathbb{1}_{\mathcal{C}_2} = 0$. Furthermore, for any $C \in \mathrm{supp}(\bar{x}^*)$ it holds that $C \subseteq S_0 \cap L$.

Let $\bar{y}^* \in [0,1]^I$ be the coverage of $\bar{x}^*$. Then $\mathrm{supp}(\bar{y}^*) \subseteq S_0 \cap L$. We note that our definition of $\bar{x}^*$ does not include the coverage of items by configurations in $T_0 \cup \mathcal{C}_2$ in $\bar{x}^0$. The coverage of these items is given by $\mathbb{1}_I - \bar{y}^*$. In the analysis we consider these coverage vectors separately, using the inequality

$$\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j}) \leq \delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j} \wedge \bar{y}^*) + \delta \sum_{j=0}^{k-1} \mathrm{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^*)\right) \ . \tag{47}$$

The configurations in $\mathrm{supp}(\bar{x}^*)$ are those that remain "difficult" after the sampling of $\mathcal{M}$; thus, $\bar{y}^*$ represents the coverage of items by these difficult configurations. Other configurations are either in T_0, or in $\mathcal{C}_2$. As the configurations in T_0 are "easy", we use them to compensate for items not selected by the matching $\mathcal{M}$. Due to a technical limitation of linear structures, we eliminate the small items from $\bar{y}^*$.

Our analysis relies on the following application of linear structures in conjunction with Lemma 2.4.

Lemma 4.8. *For $j \in \{0, 1, \ldots, k\}$, let $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ be an $\mathcal{F}_j$-measurable random vector, $\bar{w}$ the coverage of $\bar{\lambda}$, $\mathcal{S}$ an $\mathcal{F}_j$-measurable random $(\delta, \varphi(\delta))$-linear structure of $\bar{\lambda}$, and $\bar{d} \in [0,1]^I$ a small-items integral $\mathcal{F}_j$-measurable random demand vector. Then*

$$\forall r = j, \ldots, k: \quad \mathrm{OPT}_f\left(\bar{d} \wedge \mathbb{1}_{S_r}\right) \leq (1-\delta)^{r-j}(1+10\delta)\|\bar{\lambda}\| + \varphi(\delta) + \delta^{10}\mathrm{OPT}$$

with probability at least $\xi - \varphi(\delta)^2 \cdot \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right)$, where

$$\xi = \Pr\left(\forall \bar{u} \in \mathcal{S}: \ (\mathbb{1}_{S_j} \wedge \bar{d}) \cdot \bar{u} \leq \bar{w} \cdot \bar{u} + \frac{1}{\varphi^{11}(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u})\right). \tag{48}$$

The proof of the lemma is given in Section 4.1.3.

We proceed to separately bound the quantities $\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j} \wedge \bar{y}^*)$ (see Lemma 4.11) and $\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^*)\right)$ (see Lemma 4.16). The bound on $\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j} \wedge \bar{y}^*)$ is derived using the next lemmas.

Lemma 4.9. *With probability at least* $1 - \delta^{-10} \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$ *it holds that*

$$\forall h = 2, \ldots, 2 \cdot \delta^{-1}, j = 1, \ldots, k: \qquad \left| \mathbb{E}\left[\bar{x}^* \cdot \mathbb{1}_{T_{j,h}} \mid \mathcal{F}_{j-1}\right] - \bar{x}^* \cdot \mathbb{1}_{T_{j,h}} \right| \leq \delta^{20} \cdot \mathrm{OPT} \ . \tag{49}$$

The proof (given in Section 4.1.3) is a simple application of a Lemma 2.3.

Lemma 4.10. *There exists* $\mu : (0, 0.1) \to \mathbb{R}_+$*, independent of the instance* (I, v) *and* δ*, such that*

$$\forall h = 2, \ldots, 2 \cdot \delta^{-1}, j = 1, \ldots, k: \ \bar{x}^* \cdot \mathbb{1}_{U_{j,h}} \geq (1-\delta)^{h \cdot j} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - \delta^{10} \cdot \mathrm{OPT} \text{ or } \mathrm{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta) \tag{50}$$

with probability at least $1 - \delta^{-10} \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right)$.

The lemma follows from the inequality $\Pr\left(C \in U_{j,h} \mid \mathcal{F}_{j-1}\right) \geq \mathbb{1}_{C \in U_{j-1,h}} \cdot \left(1 - \frac{h}{z_j}\right)^{\alpha \cdot z_j + 1}$ implied by Lemma 2.1, the observation that $\left(1 - \frac{h}{z}\right)^{\alpha \cdot z + 1} \to (1-\delta)^h$ as $z \to \infty$, and Lemma 4.9. The dependence on μ in the lemma arises as the observation holds only if z is sufficiently large. The proof is given in Section 4.1.3. Henceforth, we use μ to denote the function in Lemma 4.10.

Lemma 4.11. *Assuming* $\mathrm{OPT} > \delta^{-30} \cdot (\varphi(\delta) + \mu(\delta))$*, with probability at least* $1 - \varphi^4(\delta) \cdot \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right)$ *it holds that*

$$\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j} \wedge \bar{y}^*) \leq \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{U_0 \setminus \mathcal{C}_2} + 30 \cdot \delta \cdot \mathrm{OPT} \ .$$

Proof. For $j = 1, \ldots, k$, define $\bar{d}^j \in [0,1]^I$, the touched demand of iteration j, as the coverage of $\bar{x}^* \wedge \mathbb{1}_{T_j}$. This is the coverage of items in configurations that become touched in iteration j, given by $\bar{d}^j_i = \sum_{C \in T_j} \bar{x}^*_C \cdot C(i)$ for all $i \in I$. For every $i \in I$ and $r \in \{0, 1, \ldots, k-1\}$ we have

$$\bar{y}^*_i - \sum_{j=1}^{r} \bar{d}^j_i = \sum_{C \in \mathcal{C}} \bar{x}^*_C \cdot C(i) - \sum_{j=1}^{r} \sum_{C \in T_j} \bar{x}^*_C \cdot C(i) = \sum_{C \in U_r} \bar{x}^*_C \cdot C(i),$$

where the last equality follows from $\mathrm{supp}(\bar{x}^*) \cap T_0 = \emptyset$ (by the definition of $\bar{x}^*$). Hence, $\bar{x}^* \wedge \mathbb{1}_{U_r}$ is a solution for LP $\left(\bar{y}^* - \sum_{j=1}^{r} \bar{d}^j\right)$, and thus $\mathrm{OPT}_f\left(\bar{y}^* - \sum_{j=1}^{r} \bar{d}^j\right) \leq \bar{x}^* \cdot \mathbb{1}_{U_r}$. It follows that for $r = 0, 1, \ldots, k-1$,

$$\begin{aligned} \mathrm{OPT}_f(\bar{y}^* \wedge \mathbb{1}_{S_r}) &\leq \sum_{j=1}^{r} \mathrm{OPT}_f\left(\bar{d}^j \wedge \mathbb{1}_{S_r}\right) + \mathrm{OPT}_f\left(\bar{y}^* - \sum_{j=1}^{r} \bar{d}^j\right) \\ &\leq \sum_{j=1}^{r} \mathrm{OPT}_f\left(\bar{d}^j \wedge \mathbb{1}_{S_r}\right) + \bar{x}^* \cdot \mathbb{1}_{U_r} \ . \end{aligned} \tag{51}$$

We use Lemma 4.8 to bound the above terms $\mathrm{OPT}_f\left(\bar{d}^j \wedge \mathbb{1}_{S_r}\right)$. We note that a natural candidate for the construction of the vector $\bar{\lambda}$ in Lemma 4.8 for iteration $j = 1, \ldots, k$ is the vector $\bar{\mu}^j \in [0,1]^{\mathcal{C}^*}$ defined by $\bar{\mu}^j_C = \sum_{C' \in T_j \text{ and } C' \cap S_j = C} \bar{x}^*_{C'}$ for all $C \in \mathcal{C}$ (and $\bar{\mu}^j_C = 0$ for $C \in \mathcal{C}^* \setminus \mathcal{C}$). It is easy to verify that $\bar{\mu}^j$ is with δ-slack and its coverage is $\bar{d}^j \wedge \mathbb{1}_{S_j}$. However, using this construction in the analysis leads to a sub-optimal approximation ratio. To some extent, this sub-optimality can be attributed to the fact that $\text{supp}(\bar{\mu}^j)$ may contain configurations which use only a small fraction of the available volume. For example, in case $C \in T_{j,h}$ for some large h and $|C \cap S_j \cap L| = 1$, we may have that $\bar{\mu}^j_{C\cap L\cap S_j} > 0$, while $v(C\cap L\cap S_j)$ is very small (e.g, $(0, 1.1\cdot\delta)$). Due to dependencies between items, such events may have non-negligible probability. To overcome this sub-optimality, we use for the construction of $\bar{\lambda}^j \in \mathbb{R}^{\mathcal{C}^*}$ conditional probabilities as described below.

For $h = 2, \ldots, 2\cdot\delta^{-1}$ and $C \in \mathcal{C}_h$, let $\bar{\gamma}^C \in [0,1]^{\mathcal{C}^*}$ be an $\frac{h}{h-1}$-relaxation of C. The existence of $\bar{\gamma}^C$ is guaranteed by Lemma 4.5. We define, for $j = 1, \ldots, k$,

$$\bar{\lambda}^j = \sum_{C\in\mathcal{C}\setminus\mathcal{C}_0} \bar{x}^*_C \cdot \left(\Pr\left(C \in T_j \mid \mathcal{F}_{j-1}\right) - \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j}\right)\cdot \mathbb{1}_{C\in U_{j-1}}\right)\cdot \bar{\gamma}^C, \tag{52}$$

and let $\bar{w}^j$ be the coverage of $\bar{\lambda}^j$. Since U_{j-1}, ρ_j and z_j are $\mathcal{F}_{j-1}$-measurable, it follows that $\bar{\lambda}^j$ is $\mathcal{F}_{j-1}$-measurable (and thus also $\mathcal{F}_j$-measurable). Furthermore, since $\bar{\gamma}^C$ is with δ-slack for every $C \in \mathcal{C}\setminus\mathcal{C}_0$, it follows that $\bar{\lambda}^j$ is with δ-slack for $j = 1, \ldots, k$.

Claim 4.12. *For $j = 1, \ldots, k$ and $i \in I$ it holds that $\mathbb{E}\left[\bar{d}^j_i \cdot \mathbb{1}_{i\in S_j} \,\middle|\, \mathcal{F}_{j-1}\right] = \bar{w}^j_i$.*

Proof. For any $i \in I \setminus L$ and $j = 1, \ldots, k$ it holds that $\mathbb{E}\left[\bar{d}^j_i \cdot \mathbb{1}_{i\in S_j} \mid \mathcal{F}_{j-1}\right] = 0 = \bar{w}^j_i$, as $\text{supp}(\bar{y}^*) \subseteq L$ and $\bar{y}^*$ is the coverage of $\bar{x}^*$. Thus, it remains to handle the case in which $i \in L$.

Now, for every $i \in L$ and $j = 1, \ldots, k$, we have

$$\begin{aligned}
\mathbb{E}\left[\bar{d}^j_i \cdot \mathbb{1}_{i\in S_j} \,\middle|\, \mathcal{F}_{j-1}\right] &= \mathbb{E}\left[\sum_{C\in\mathcal{C}} \mathbb{1}_{C\in T_j}\cdot \mathbb{1}_{i\in S_j}\cdot \bar{x}^*_C \cdot C(i) \,\middle|\, \mathcal{F}_{j-1}\right] \\
&= \mathbb{E}\left[\sum_{C\in\mathcal{C}\setminus\mathcal{C}_0} \left(\mathbb{1}_{C\in T_j} - \mathbb{1}_{C\in T_j}\cdot \mathbb{1}_{i\notin S_j}\right)\cdot \bar{x}^*_C \cdot C(i) \,\middle|\, \mathcal{F}_{j-1}\right] \\
&= \sum_{C\in\mathcal{C}\setminus\mathcal{C}_0} \left(\Pr\left(C\in T_j \mid \mathcal{F}_{j-1}\right) - \mathbb{E}\left[\mathbb{1}_{i\notin S_j}\mathbb{1}_{C\in U_{j-1}} \mid \mathcal{F}_{j-1}\right]\right)\cdot \bar{x}^*_C \cdot C(i)\ .
\end{aligned} \tag{53}$$

The second equality uses $T_j \cap \mathcal{C}_0 = \emptyset$ for $j \geq 1$, and the third equality uses that

$$\mathbb{1}_{C\in T_j}\mathbb{1}_{i\notin S_j} = \mathbb{1}_{C\in U_{j-1}}\cdot \mathbb{1}_{C\notin U_j}\cdot \mathbb{1}_{i\notin S_j} = \mathbb{1}_{C\in U_{j-1}}\cdot \mathbb{1}_{i\notin S_j}$$

for any configuration C for which $i \in C$. By Lemma 2.1, we have

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}_{i\notin S_j}\mathbb{1}_{C\in U_{j-1}} \,\middle|\mathcal{F}_{j-1}\right] &= \mathbb{1}_{C\in U_{j-1}}\cdot \mathbb{E}\left[\mathbb{1}_{i\notin S_j} \,\middle|\mathcal{F}_{j-1}\right] \\
&= \mathbb{1}_{C\in U_{j-1}}\left(1 - \mathbb{1}_{i\in S_{j-1}}\left(1-\frac{1}{z_j}\right)^{\rho_j}\right) \\
&= \mathbb{1}_{C\in U_{j-1}}\left(1 - \left(1-\frac{1}{z_j}\right)^{\rho_j}\right)
\end{aligned}$$

for any $C \in \mathcal{C}\setminus\mathcal{C}_0$ and $i \in C \cap L$. Furthermore, since $\bar{\gamma}^C$ is a relaxation of C, we have that

$C(i) = \sum_{C' \in \mathcal{C}^*} \bar{\gamma}^C_{C'} \cdot C'(i)$. Therefore, for any $C \in \mathcal{C} \setminus \mathcal{C}_0$ and $i \in L$, it holds that

$$\begin{aligned}
&\left(\Pr\left(C \in T_j \mid \mathcal{F}_{j-1}\right) - \mathbb{E}\left[\mathbb{1}_{i \notin S_j} \mathbb{1}_{C \in U_{j-1}} \mid \mathcal{F}_{j-1}\right]\right) \cdot \bar{x}^*_C \cdot C(i) \\
&\quad = \left(\Pr\left(C \in T_j \mid \mathcal{F}_{j-1}\right) - \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j}\right) \cdot \mathbb{1}_{C \in U_{j-1}}\right) \cdot \bar{x}^*_C \cdot C(i) \\
&\quad = \left(\Pr\left(C \in T_j \mid \mathcal{F}_{j-1}\right) - \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j}\right) \cdot \mathbb{1}_{C \in U_{j-1}}\right) \cdot \bar{x}^*_C \cdot \sum_{C' \in \mathcal{C}^*} \bar{\gamma}^C_{C'} \cdot C'(i) \ .
\end{aligned} \tag{54}$$

By incorporating (54) into (53), we have (for every $i \in L$ and $j = 1, \ldots, k$) that

$$\begin{aligned}
&\mathbb{E}\left[\bar{d}^j_i \cdot \mathbb{1}_{i \in S_j} \;\middle|\; \mathcal{F}_{j-1}\right] \\
&= \sum_{C \in \mathcal{C} \setminus \mathcal{C}_0} \bar{x}^*_C \cdot \left(\Pr\left(C \in T_j \mid \mathcal{F}_{j-1}\right) - \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j}\right) \cdot \mathbb{1}_{C \in U_{j-1}}\right) \cdot \sum_{C' \in \mathcal{C}^*} \bar{\gamma}^C_{C'} \cdot C'(i) \\
&= \sum_{C' \in \mathcal{C}^*} C'(i) \cdot \sum_{C \in \mathcal{C} \setminus \mathcal{C}_0} \bar{x}^*_C \cdot \left(\Pr\left(C \in T_j \mid \mathcal{F}_{j-1}\right) - \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j}\right) \cdot \mathbb{1}_{C \in U_{j-1}}\right) \cdot \bar{\gamma}^C_{C'} \\
&= \sum_{C' \in \mathcal{C}^*} C'(i) \cdot \bar{\lambda}^j_{C'} = \bar{w}^j_i \ .
\end{aligned}$$

$\diamond$

To show the existence of a linear structure for $\bar{\lambda}$ using Lemma 4.2, we also need the following claim.

Claim 4.13. *For $j = 1, \ldots, k$ it holds that $\bar{\lambda}^j \in [0,1]^{\mathcal{C}^*}$, $\bar{w}^j \in [0,1]^I$, and $\bar{\lambda}^j$ is small-items integral.*

Proof. We first show that $\bar{\lambda}^j \in \mathbb{R}^{\mathcal{C}^*}_{\geq 0}$. Let $C \in \mathcal{C} \setminus \mathcal{C}_0$, thus there is $i \in C \cap L$. It therefore holds that

$$\begin{aligned}
\Pr(C \in T_j \mid \mathcal{F}_{j-1}) &= \mathbb{E}\left[\mathbb{1}_{C \in U_{j-1}} \cdot \mathbb{1}_{C \notin U_j} \mid \mathcal{F}_{j-1}\right] \\
&= \mathbb{1}_{C \in U_{j-1}} \cdot \Pr(C \notin U_j \mid \mathcal{F}_{j-1}) \\
&\geq \mathbb{1}_{C \in U_{j-1}} \cdot \Pr(i \notin S_j \mid \mathcal{F}_{j-1}) \\
&= \mathbb{1}_{C \in U_{j-1}} \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j} \cdot \mathbb{1}_{i \in S_{j-1}}\right) \\
&= \mathbb{1}_{C \in U_{j-1}} \left(1 - \left(1 - \frac{1}{z_j}\right)^{\rho_j}\right) \ .
\end{aligned} \tag{55}$$

The inequality holds since $i \notin S_j$ implies $C \notin U_j$, the third equality is by Lemma 2.1, and the last equality holds since $\mathbb{1}_{C \in U_{j-1}} \cdot \mathbb{1}_{i \in S_{j-1}} = \mathbb{1}_{C \in U_{j-1}}$. By (55) it follows that $\bar{\lambda}^j \in \mathbb{R}^{\mathcal{C}^*}_{\geq 0}$.

Since $\bar{w}^j_i = \mathbb{E}\left[\bar{d}^j_i \cdot \mathbb{1}_{i \in S_j} \mid \mathcal{F}_{j-1}\right] \leq 1$ for every $i \in I$ for $j = 1, \ldots, k$ (Claim 4.12) it follows that $\bar{w}^j \in [0,1]^I$ and subsequently $\bar{\lambda}^j \in [0,1]^{\mathcal{C}^*}$ for $j = 1, \ldots, k$. Furthermore, $\bar{w}^j_i = 0$ for every $i \in I \setminus L$ (as $\bar{y}^*_i = 0$, $\bar{y}^*$ is the coverage of $\bar{x}^*$ and (52)), hence $\bar{w}^j$ and $\bar{\lambda}^j$ are small-items integral. $\diamond$

By Lemma 4.2 there is a $(\delta, \varphi(\delta))$-linear structure $\mathcal{S}_j$ of $\bar{\lambda}^j$ for $j = 1, \ldots, k$.

Claim 4.14. *For any $j \in \{1, \ldots, k\}$ it holds that*

$$\Pr\left(\forall \bar{u} \in \mathcal{S}_j : \ (\mathbb{1}_{S_j} \wedge \bar{d}^j) \cdot \bar{u} \ \leq \ \mathbb{E}\left[(\mathbb{1}_{S_j} \wedge \bar{d}^j) \cdot \bar{u} \mid \mathcal{F}_{j-1}\right] + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)} \cdot \mathsf{tol}(\bar{u})\right) \geq 1 - \varphi(\delta) \cdot \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \ .$$

The proof of Claim 4.14, given in Section 4.1.3, follows from Lemma 2.3. By Claim 4.12 it holds that $\mathbb{E}\left[\bar{u}\cdot\left(\bar{d}^j \wedge \mathbb{1}_{S_j}\right)\middle|\mathcal{F}_{j-1}\right] = \bar{u}\cdot\bar{w}^j$ for $j=1,\dots,k$ and $\bar{u}\in\mathcal{S}_j$; therefore,

$$\begin{aligned}
&\Pr\left(\forall \bar{u}\in\mathcal{S}_j:\ (\mathbb{1}_{S_j}\wedge\bar{d}^j)\cdot\bar{u} \le \bar{w}^j\cdot\bar{u} + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\cdot\mathsf{tol}(\bar{u})\right)\\
&\quad= \Pr\left(\forall \bar{u}\in\mathcal{S}_j:\ (\mathbb{1}_{S_j}\wedge\bar{d}^j)\cdot\bar{u} \le \mathbb{E}\left[(\mathbb{1}_{S_j}\wedge\bar{d}^j)\cdot\bar{u}\middle|\mathcal{F}_{j-1}\right] + \frac{\mathrm{OPT}\cdot\mathsf{tol}(\bar{u})}{\varphi^{11}(\delta)}\right)\\
&\quad\ge 1-\varphi(\delta)\cdot\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \ .
\end{aligned}$$

Here, the last inequality follows from Claim 4.14. Thus, by Lemma 4.8, with probability at least

$$1-k\cdot\varphi(\delta)\cdot\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) - k\cdot\varphi^2(\delta)\cdot\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \ge 1-\varphi^3(\delta)\cdot\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right),$$

it holds that

$$\forall j=1,\dots,k, r=j,\dots,k:\ \mathrm{OPT}_f\left(\bar{d}^j\wedge\mathbb{1}_{S_r}\right) \le (1-\delta)^{r-j}(1+10\delta)\|\bar{\lambda}^j\| + \varphi(\delta) + \delta^{10}\mathrm{OPT} \ . \tag{56}$$

We henceforth assume that (56), (49) and (50) hold.

Observe that, for $j=1,\dots,k$,

$$\begin{aligned}
\|\bar{\lambda}^j\| &= \sum_{C\in\mathcal{C}\setminus\mathcal{C}_0} \bar{x}^*_C\cdot\left(\Pr\left(C\in T_j \mid \mathcal{F}_{j-1}\right) - \left(1-\left(1-\frac{1}{z_j}\right)^{\rho_j}\right)\cdot\mathbb{1}_{C\in U_{j-1}}\right)\cdot\|\bar{\gamma}^C\|\\
&\le \sum_{C\in\mathcal{C}\setminus\mathcal{C}_0} \bar{x}^*_C\cdot\left(\Pr\left(C\in T_j \mid \mathcal{F}_{j-1}\right) - \delta\cdot\mathbb{1}_{C\in U_{j-1}}\right)\cdot\|\bar{\gamma}^C\|\\
&\le \sum_{h=2}^{2\cdot\delta^{-1}}\sum_{C\in\mathcal{C}_h} \bar{x}^*_C\cdot\left(\Pr\left(C\in T_j \mid \mathcal{F}_{j-1}\right) - \delta\cdot\mathbb{1}_{C\in U_{j-1}}\right)\cdot\frac{h}{h-1}\\
&= \sum_{h=2}^{2\cdot\delta^{-1}}\frac{h}{h-1}\left(\mathbb{E}\left[\bar{x}^*\cdot\mathbb{1}_{T_{j,h}}\ \middle|\ \mathcal{F}_{j-1}\right] - \delta\cdot\bar{x}^*\cdot\mathbb{1}_{U_{j-1,h}}\right)\\
&\le \sum_{h=2}^{2\cdot\delta^{-1}}\frac{h}{h-1}\left(\bar{x}^*\cdot\mathbb{1}_{T_{j,h}} + \delta^{10}\cdot\mathrm{OPT} - \delta\cdot\bar{x}^*\cdot\mathbb{1}_{U_{j-1,h}}\right)\\
&\le \sum_{h=2}^{2\cdot\delta^{-1}}\frac{h}{h-1}\left((1-\delta)\bar{x}^*\cdot\mathbb{1}_{U_{j-1,h}} - \bar{x}^*\cdot\mathbb{1}_{U_{j,h}}\right) + \delta^8\cdot\mathrm{OPT} \ .
\end{aligned}\tag{57}$$

The first inequality follows from $\left(1-\frac{1}{z_j}\right)^{\rho_j}\le(1-\delta)$ (Lemma 2.1). The second inequality holds, since $\bar{\gamma}^C$ is an $\frac{h}{h-1}$-relaxation of C for any $C\in\mathcal{C}_h$; the third inequality follows from the assumption that (49) holds; and the last inequality uses $T_{j,h}=U_{j-1,h}\setminus U_{j,h}$.

Combining (56) and (57) with $\mathrm{OPT}>\delta^{-30}\varphi(\delta)$, we have

$$\frac{\mathrm{OPT}_f\left(\bar{d}^j\wedge\mathbb{1}_{S_r}\right)}{1+10\delta} \le (1-\delta)^{r-j}\sum_{h=2}^{2\cdot\delta^{-1}}\frac{h}{h-1}\left((1-\delta)\bar{x}^*\cdot\mathbb{1}_{U_{j-1,h}} - \bar{x}^*\cdot\mathbb{1}_{U_{j,h}}\right) + \delta^7\mathrm{OPT}$$

for $j = 1, \ldots, k$ and $r = j, \ldots, k$. Using the last inequality and (51), we obtain

$$\begin{aligned}
\frac{\mathrm{OPT}_f(\bar{y}^* \wedge \mathbb{1}_{S_r})}{1+10\delta} &\leq \sum_{j=1}^{r}(1-\delta)^{r-j} \sum_{h=2}^{2\cdot\delta^{-1}} \frac{h}{h-1}\left((1-\delta)\bar{x}^* \cdot \mathbb{1}_{U_{j-1,h}} - \bar{x}^* \cdot \mathbb{1}_{U_{j,h}}\right) + \bar{x}^* \cdot \mathbb{1}_{U_r} + \delta^5 \mathrm{OPT} \\
&= \sum_{h=2}^{2\cdot\delta^{-1}} \frac{h}{h-1}\left((1-\delta)^r \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - \bar{x}^* \cdot \mathbb{1}_{U_{r,h}}\right) + \bar{x}^* \cdot \mathbb{1}_{U_r} + \delta^5 \mathrm{OPT} \\
&= \sum_{h=2}^{2\cdot\delta^{-1}} \frac{1}{h-1}\left((1-\delta)^r \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - \bar{x}^* \cdot \mathbb{1}_{U_{r,h}}\right) + (1-\delta)^r \bar{x}^* \cdot \mathbb{1}_{U_0} + \delta^5 \mathrm{OPT}
\end{aligned}$$

for every $r \in \{0, 1, \ldots, k-1\}$. Observe that $\mathrm{OPT}_f(\bar{y}^* \wedge \mathbb{1}_{S_r}) \leq \mathrm{OPT}_f(\mathbb{1}_{S_r}) \leq \mathrm{OPT}_f(\mathbb{1}_{S_j})$ for $j = 1, \ldots, k$ and $r = j, \ldots, k$; thus, if $\mathrm{OPT}(\mathbb{1}_{S_j}) \leq \mu(\delta) \leq \delta^{30}\mathrm{OPT}$ for some $j \in \{1, \ldots, k\}$, then for every $r \geq j$ it holds that $\mathrm{OPT}_f(\bar{y}^* \wedge \mathbb{1}_{S_r}) \leq \delta^{30}\mathrm{OPT}$. Using the above inequality and (50), we have

$$\frac{\mathrm{OPT}_f(\bar{y}^* \wedge \mathbb{1}_{S_r})}{1+10\delta} \leq \sum_{h=2}^{2\cdot\delta^{-1}} \frac{(1-\delta)^r - (1-\delta)^{h\cdot r}}{h-1} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} + (1-\delta)^r \bar{x}^* \cdot \mathbb{1}_{U_0} + \delta^4 \mathrm{OPT} \ .$$

Thus,

$$\begin{aligned}
&\frac{\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j} \wedge \bar{y}^*)}{1+10\delta} \\
\leq\ & \delta \sum_{j=0}^{k-1} \sum_{h=2}^{2\cdot\delta^{-1}} \frac{(1-\delta)^j - (1-\delta)^{h\cdot j}}{h-1} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} + \delta \cdot \sum_{j=0}^{k-1}(1-\delta)^j \bar{x}^* \cdot \mathbb{1}_{U_0} + \delta^3 \mathrm{OPT} \\
=\ & \delta \sum_{h=2}^{2\cdot\delta^{-2}} \frac{\bar{x}^* \cdot \mathbb{1}_{U_{0,h}}}{h-1}\left(\frac{1-(1-\delta)^k}{1-(1-\delta)} - \frac{1-(1-\delta)^{k\cdot h}}{1-(1-\delta)^h}\right) + \delta \cdot \frac{1-(1-\delta)^k}{1-(1-\delta)} \cdot \bar{x}^* \cdot \mathbb{1}_{U_0} + \delta^3 \mathrm{OPT} \\
\leq\ & \sum_{h=2}^{2\cdot\delta^{-2}} \frac{\bar{x}^* \cdot \mathbb{1}_{U_{0,h}}}{h-1}\left(1 - \frac{1-\delta}{h}\right) + \bar{x}^* \cdot \mathbb{1}_{U_0} + \delta^3 \mathrm{OPT} \\
\leq\ & \sum_{h=3}^{2\cdot\delta^{-2}} \frac{h+1}{h} \cdot \bar{x}^0 \cdot \mathbb{1}_{U_{0,h}} + \delta^3 \mathrm{OPT} + \delta \|\bar{x}^*\| \ .
\end{aligned}$$

The second inequality holds, since $(1-\delta)^k \leq \delta$ and $(1-\delta)^h \geq 1 - \delta h$. The last inequality uses $\bar{x}^* \cdot \mathbb{1}_{U_{0,h}} = \bar{x}^0 \cdot \mathbb{1}_{U_{0,h}}$ for $h \geq 3$, and $\bar{x}^* \cdot \mathbb{1}_{\mathcal{C}_2} = 0$ by the definition of $\bar{x}^*$. Since $\|\bar{x}^*\| \leq \|\bar{x}^0\| \leq (1+\delta^2)\mathrm{OPT} \leq 1.01 \cdot \mathrm{OPT}$, we have

$$\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j} \wedge \bar{y}^*) \leq \sum_{h=3}^{2\cdot\delta^{-2}} \frac{h+1}{h} \cdot \bar{x}^0 \cdot \mathbb{1}_{U_{0,h}} + 30 \cdot \delta \cdot \mathrm{OPT} \leq \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{U_0 \setminus \mathcal{C}_2} + 30\delta \cdot \mathrm{OPT}, \quad (58)$$

as in the statement of the lemma. As we assumed that (56), (49) and (50) hold, by Lemma 4.9 and Lemma 4.10 it follows that (58) holds with probability at least

$$1 - \varphi^3(\delta) \cdot \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) - 2 \cdot \delta^{-10} \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right) \geq 1 - \varphi^4(\delta) \cdot \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \ . \qquad \square$$

Define $\bar{y}^{\mathcal{M}}$ as the coverage of $\bar{x}^0 \wedge \mathbb{1}_{\mathcal{C}_2}$; that is, $\bar{y}^{\mathcal{M}}_i = \sum_{C \in \mathcal{C}_2} \bar{x}^0_C \cdot C(i)$ for all $i \in I$. To obtain a bound on $\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^0)\right)$, we use the next lemma.

Lemma 4.15. *For any $i \in I$ it holds that $\Pr(i \notin S_0) = (1-\delta^4)\bar{y}^{\mathcal{M}}_i$ if $i \in L$, and $\Pr(i \notin S_0) = 0$ otherwise.*

Proof. Let $G = (L, E)$ be the δ-matching graph of the instance. We use $N(i)$ to denote the set of neighbors of $i \in L$. Since $\mathcal{M}$ is a matching, for every $i \in L$ it holds that $\mathbb{1}_{i \notin S_0} = \sum_{i' \in N(i)} \mathbb{1}_{\{i,i'\} \in \mathcal{M}}$. Therefore, for any $i \in L$ it holds that

$$\begin{aligned} \Pr(i \notin S_0) = \mathbb{E}[\mathbb{1}_{i \notin S_0}] = \sum_{i' \in N(i)} \mathbb{E}\left[\mathbb{1}_{\{i,i'\} \in \mathcal{M}}\right] &= (1-\delta^4) \sum_{i' \in N(i)} \sum_{C \in \mathcal{C}_2 \text{ s.t. } \{i,i'\} \subseteq C} \bar{x}^0_C \\ &= (1-\delta^4) \sum_{C \in \mathcal{C}_2} \bar{x}^0_C \cdot C(i) = (1-\delta^4) \cdot \bar{y}^{\mathcal{M}}_i \ . \end{aligned}$$

The third equality holds, since $\Pr(e \in \mathcal{M}) = (1-\delta^4) \sum_{C \in \mathcal{C}_2 \text{ s.t. } e \subseteq C} \bar{x}^0_C$. Also, for any $i \in I \setminus L$ it holds that $i \notin \bigcup_{e \in \mathcal{M}} e$; thus, $i \in S_0$, i.e., $\Pr(i \notin S_0) = 0$. $\square$

We now derive an upper bound for $\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^*)\right)$.

Lemma 4.16. *Assuming $\mathrm{OPT} > \delta^{-30}\varphi(\delta)$, with probability at least $1 - \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)} + \varphi^2(\delta) \cdot \ln \mathrm{OPT}\right)$ it holds that*

$$\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^*)\right) \le \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{T_0 \setminus \mathcal{C}_2} + \frac{1}{3} \cdot |\mathcal{M}| + 50\delta \cdot (\mathrm{OPT} + |\mathcal{M}|) \ .$$

Proof. Similar to the proof of Lemma 4.11, we use Lemma 4.8 also in this proof. To this end, we construct a vector $\bar{\lambda}$ that is used to derive a linear structure $\mathcal{S}$. Subsequently, we show that $\bar{\lambda}$ and $\mathcal{S}$ admit the conditions of Lemma 4.8 with respect to the demand vector $\mathbb{1}_{S_0} \wedge (\mathbb{1}_I - \bar{y}^*)$.

For any $h = 2, \ldots, 2 \cdot \delta^{-1}$ and $C \in \mathcal{C}_h$, let $\bar{\gamma}^C$ be an $\frac{h}{h-1}$-relaxation of C, and for any $C \in \mathcal{C}_0$ let $\bar{\gamma}^C$ be a $(1+4\delta)$-relaxation of C. Furthermore, for any $C \in \mathcal{C}$ such that $v(C) \le (\delta, \delta)$ let $\bar{\tau}^C$ be a 4δ-relaxation of C. The existence of these relaxations is guaranteed by Lemma 4.4, Lemma 4.5, and Lemma 4.6. Define

$$\bar{\lambda} = \delta^4 \sum_{i \in L} \bar{y}^{\mathcal{M}}_i \cdot \mathbb{1}_{\{\{i\}\}} + \sum_{C \in T_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot \bar{\gamma}^C + \sum_{C \in U_0 \cup \mathcal{C}_2} \bar{x}^0_C \cdot \bar{\tau}^{C \setminus L},$$

where $\mathbb{1}_{\{\{i\}\}} \in [0,1]^{\mathcal{C}^*} = \bar{z}$ such that $\bar{z}_{\{i\}} = 1$, and $\bar{z}_C = 0$ for $C \in \mathcal{C}^* \setminus \{\{i\}\}$. Observe that $\mathcal{C}_0 \subseteq T_0$ by definition; thus, $v(C \setminus L) \le (\delta, \delta)$ for every $C \in U_0 \cup \mathcal{C}_2$. That is, $\bar{\lambda}$ is well-defined. Since the instance does not contain δ-huge items, it follows that $\mathbb{1}_{\{\{i\}\}}$ is with δ-slack. Hence, $\bar{\lambda}$ is with δ-slack as well. As T_0 and U_0 are $\mathcal{F}_0$-measurable, it follows that $\bar{\lambda}$ is $\mathcal{F}_0$-measurable. Let $\bar{w}$ be the coverage of $\bar{\lambda}$ and define $\bar{d} = \mathbb{1}_{S_0} \wedge (1 - \bar{y}^*)$. Observe that we may have $\bar{w}_i > 0$ (i.e., $i \in \mathrm{supp}(\bar{w})$) for items already selected by the matching, that is, items in $L \setminus S_0$. The coverage of these items can intuitively be viewed as a placeholder for items $i \in L \cap S_0$ for which $\bar{w}_i < \bar{d}_i$.

For any $i \in I \setminus L$, it holds that

$$\begin{aligned} \bar{w}_i = \sum_{C \in \mathcal{C}^*} \bar{\lambda}_C \cdot C(i) &= \sum_{C \in T_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) + \sum_{C \in U_0 \cup \mathcal{C}_2} \bar{x}^0_C \cdot C(i) \\ &= \sum_{C \in \mathcal{C}} \bar{x}^0_C \cdot C(i) = 1 = \mathbb{1}_{i \in S_0}(1 - \bar{y}^*_i) = \bar{d}_i \ . \end{aligned} \tag{59}$$

The fourth equality holds, as $\bar{x}^0$ is a solution for MLP. The fifth equality holds, since $\bar{y}^*_i = 0$ for all $i \in I \setminus L$ and by Lemma 4.15. In particular, it follows that $\bar{w}$ and $\bar{\lambda}$ are small-items integral, and $\bar{w}_i - \bar{d}_i = 0$ for any $i \in I \setminus L$. Furthermore, for any $i \in L$ it holds that

$$\bar{w}_i = \delta^4 \cdot \bar{y}^{\mathcal{M}}_i + \sum_{C \in T_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) \leq \delta^4 \cdot \bar{y}^{\mathcal{M}}_i + \sum_{C \in \mathcal{C} \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) = \delta^4 \cdot \bar{y}^{\mathcal{M}}_i + (1 - \bar{y}^{\mathcal{M}}_i) \leq 1,$$

thus $\bar{w} \in [0,1]^I$ and we can infer that $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$.

For any $i \in L$, we have

$$\begin{aligned}
\bar{d}_i &= \mathbb{1}_{i \in S_0} \left(1 - \sum_{C \in U_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) \right) \\
&= \mathbb{1}_{i \in S_0} - (1 - \mathbb{1}_{i \notin S_0}) \sum_{C \in U_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) \\
&= \mathbb{1}_{i \in S_0} - \sum_{C \in U_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) - \sum_{C \in \mathcal{C} \setminus \mathcal{C}_2} \mathbb{1}_{i \notin S_0} \cdot \mathbb{1}_{C \in U_0} \cdot \bar{x}^0_C \cdot C(i) \\
&= \mathbb{1}_{i \in S_0} - \sum_{C \in U_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i),
\end{aligned}$$

where the the fourth equality holds since for every $C \in \mathcal{C}$ such that $i \in C$, if $i \notin S_0$ then $C \notin U_0$. Thus, for every $i \in L$,

$$\begin{aligned}
\bar{w}_i - \bar{d}_i &= \delta^4 \cdot \bar{y}^{\mathcal{M}}_i + \sum_{C \in T_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) - \left(\mathbb{1}_{i \in S_0} - \sum_{C \in U_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) \right) \\
&= \delta^4 \cdot \bar{y}^{\mathcal{M}}_i + \sum_{C \in \mathcal{C} \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) - \mathbb{1}_{i \in S_0} \\
&= \delta^4 \cdot \bar{y}^{\mathcal{M}}_i + 1 - \bar{y}^{\mathcal{M}}_i - \mathbb{1}_{i \in S_0} \\
&= \mathbb{1}_{i \notin S_0} - (1 - \delta^4) \cdot \bar{y}^{\mathcal{M}}_i,
\end{aligned} \tag{60}$$

where the third equality holds since

$$1 = \sum_{C \in \mathcal{C}} \bar{x}^0_C \cdot C(i) = \sum_{C \in \mathcal{C} \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) + \sum_{C \in \mathcal{C}_2} \bar{x}^0_C \cdot C(i) = \sum_{C \in \mathcal{C} \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) + \bar{y}^{\mathcal{M}}_i \ .$$

By (59), (60) and Lemma 4.15, it holds that $\mathbb{E}[\bar{w}_i] = \mathbb{E}[\bar{d}_i]$ for every $i \in I$.

Using the concentration bounds for SampleMatching, as given by Chekuri et al. [CVZ11], we can show that, with high probability, $\bar{u} \cdot \bar{d} \lesssim \bar{u} \cdot \bar{w}$ for every $\bar{u} \in \mathbb{R}^I_{\geq 0}$.

Claim 4.17. *For any $\bar{u} \in \mathbb{R}^I_{\geq 0}$ it holds that*

$$\Pr\left(\bar{d} \cdot \bar{u} > \bar{w} \cdot \bar{u} + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)} \cdot \mathsf{tol}(\bar{u}) \right) \leq \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)} \right) \ .$$

The proof of Claim 4.17 is given in Section 4.1.3.

Let $\mathcal{S}^* \subseteq \mathbb{R}^I_{\geq 0}$ be the set defined in Lemma 4.2. Also, by Lemma 4.2, there exists a $(\delta, \varphi(\delta))$-linear structure $\mathcal{S}$ of $\bar{\lambda}$ such that for any $\bar{u} \in \mathcal{S}$ which satisfies $\mathrm{supp}(\bar{u}) \cap L \neq \emptyset$ it holds that $\bar{u} \in \mathcal{S}^*$. Observe that $\mathcal{S}^*$ is non-random while $\mathcal{S}$ is an $\mathcal{F}_0$-measurable random set, as $\bar{\lambda}$ is $\mathcal{F}_0$-measurable.

Claim 4.17 requires that the vector $\bar{u} \in \mathbb{R}^I_{\geq 0}$ is deterministic, and thus we cannot directly use the claim with a random vector $\bar{u} \in \mathcal{S}$. Instead, we use the set $\mathcal{S}^*$ to circumvent this issue. Observe that for any $\bar{u} \in \mathcal{S}$, if $\text{supp}(\bar{u}) \cap L = \emptyset$ then $\bar{d} \cdot \bar{u} = \bar{w} \cdot \bar{w}$ by (59), and if $\text{supp}(\bar{u}) \neq \emptyset$ then $\bar{u} \in \mathcal{S}^*$. Thus,

$$\Pr\left(\forall \bar{u} \in \mathcal{S}:\ \bar{d} \cdot \bar{u} \leq \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{11}(\delta)} \cdot \mathsf{tol}(\bar{u})\right) \geq \Pr\left(\forall \bar{u} \in \mathcal{S}^*:\ \bar{d} \cdot \bar{u} \leq \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{11}(\delta)} \cdot \mathsf{tol}(\bar{u})\right)$$
$$\geq 1 - |\mathcal{S}^*| \cdot \exp\left(-\frac{\text{OPT}}{\varphi^{25}(\delta)}\right) \geq 1 - \exp\left(-\frac{\text{OPT}}{\varphi^{25}(\delta)} + \varphi(\delta) \cdot \ln \text{OPT}\right) .$$

The second inequality is by the union bound, and Claim 4.17. The third inequality holds, since $|\mathcal{S}^*| \leq \varphi(\delta) \cdot |L|^4 \leq \varphi(\delta) \cdot 2^4 \cdot \delta^{-4} \cdot \text{OPT}^4$ as $\text{OPT} \geq \frac{\delta}{2}|L|$. Therefore, by Lemma 4.8, it holds that

$$\forall j = 0, \ldots, k: \quad \text{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^*)\right) \leq (1-\delta)^j \cdot (1+10\delta)\|\bar{\lambda}\| + \varphi(\delta) + \delta^{10}\text{OPT} \tag{61}$$

with probability at least

$$1-\exp\left(-\frac{\text{OPT}}{\varphi^{25}(\delta)} + \varphi(\delta) \cdot \ln \text{OPT}\right) - \varphi^2(\delta) \cdot \exp\left(-\frac{\text{OPT}}{\varphi^{25}(\delta)}\right) \geq 1 - \exp\left(-\frac{\text{OPT}}{\varphi^{25}(\delta)} + \varphi^2(\delta) \cdot \ln \text{OPT}\right) .$$

We henceforth assume that (61) holds.

We note that

$$\begin{aligned}\|\bar{\lambda}\| &\leq \delta^4 \cdot \mathbb{1}_L \cdot \bar{y}^{\mathcal{M}} + \sum_{h=3}^{2\cdot\delta^{-1}} \frac{h}{h-1} \cdot \bar{x}^0 \cdot \mathbb{1}_{T_{0,h}} + (1+4\delta) \cdot \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C}_0} + 4\delta\|\bar{x}^0\| \\ &\leq \frac{4}{3} \cdot \mathbb{1}_{T_0 \setminus \mathcal{C}_2} \cdot \bar{x}^0 + \frac{1}{6} \cdot \sum_{h=3}^{2\cdot\delta^{-1}} \bar{x}^0 \cdot \mathbb{1}_{T_{0,h}} + 10\delta \cdot \text{OPT},\end{aligned} \tag{62}$$

where the second inequality uses

$$\mathbb{1}_L \cdot \bar{y}^{\mathcal{M}} = \sum_{i \in L} \bar{y}^{\mathcal{M}}_i = \sum_{i \in L} \sum_{C \in \mathcal{C}_2} \bar{x}^0_C \cdot C(i) = \sum_{C \in \mathcal{C}_2} \bar{x}^0_C \cdot 2 \leq 2 \cdot \bar{x}^0 \cdot \mathbb{1}_{\mathcal{C}_2} \leq 2 \cdot (1+\delta^2)\text{OPT} .$$

It also holds that

$$\begin{aligned}\sum_{h=3}^{2\cdot\delta^{-1}} \bar{x}^0 \cdot \mathbb{1}_{T_{0,h}} &= \sum_{C \in \mathcal{C} \setminus \mathcal{C}_0 \setminus \mathcal{C}_2} \bar{x}^0_C \cdot \mathbb{1}_{C \in T_0} \\ &\leq \sum_{C \in \mathcal{C} \setminus \mathcal{C}_0 \setminus \mathcal{C}_2} \bar{x}^0_C \sum_{i \in C \cap L} \mathbb{1}_{i \notin S_0} \leq \sum_{i \in L} \mathbb{1}_{i \notin S_0} \sum_{C \in \mathcal{C} \setminus \mathcal{C}_2} \bar{x}^0_C \cdot C(i) \leq \sum_{i \in L} \mathbb{1}_{i \notin S_0} \leq 2 \cdot |\mathcal{M}| .\end{aligned}$$

Plugging the above inequality into (62), we obtain

$$\|\bar{\lambda}\| \leq \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{T_0 \setminus \mathcal{C}_2} + \frac{1}{3} \cdot |\mathcal{M}| + 10 \cdot \delta \cdot \text{OPT} . \tag{63}$$

By (61) and (63), we have

$$\begin{aligned}\delta \sum_{j=0}^{k-1} \text{OPT}_f\left(\mathbb{1}_{S_j} \wedge (\mathbb{1}_I - \bar{y}^*)\right) &\leq \delta \sum_{j=0}^{k-1} \left((1-\delta)^j \cdot (1+10\delta)\|\bar{\lambda}\| + \varphi(\delta) + \delta^{10}\text{OPT}\right) \\ &\leq (1+10\delta)\|\bar{\lambda}\| + \delta^8\text{OPT} \\ &\leq (1+10\delta)\left(\frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{T_0 \setminus \mathcal{C}_2} + \frac{1}{3} \cdot |\mathcal{M}| + 10\delta \cdot \text{OPT}\right) + \delta^8\text{OPT} \\ &\leq \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{T_0 \setminus \mathcal{C}_2} + \frac{1}{3} \cdot |\mathcal{M}| + 50\delta(\text{OPT} + |\mathcal{M}|),\end{aligned} \tag{64}$$

where the second inequality uses $\mathrm{OPT} > \delta^{-30}\varphi(\delta)$, and the last inequality holds since $\|\bar{x}^0\| \leq 1.01 \cdot \mathrm{OPT}$. As we assumed that (61) holds, it follows that inequality (64) holds with probability at least $1 - \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)} + \varphi^2(\delta) \cdot \ln \mathrm{OPT}\right)$, as stated in the lemma. $\square$

4.1.2 Asymptotic Approximation Ratio

Proof of Lemma 1.9. Note that we may assume OPT is larger than any function which depends on δ (but not on the instance). Assume that the statements of Lemmas 2.6, 4.7, 4.11 and 4.16 hold. This occurs with probability at least

$$1-\delta^{-2}\cdot\exp(-\delta^7\cdot\mathrm{OPT})-\exp(-\delta^{10}\cdot\mathrm{OPT})-\varphi^4(\delta)\cdot\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right)-\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}+\varphi^2(\delta)\cdot\ln\mathrm{OPT}\right) \geq \frac{1}{2},$$

assuming that OPT is sufficiently large.

We also assume that $\mathrm{OPT} > \delta^{-30}\left(\varphi(\delta) + \mu(\delta)\right)$. By Lemmas 4.11 and 4.16, we have

$$\begin{aligned}
\sum_{j=1}^{k} \rho_j &\leq k + (1+2\delta)\delta \sum_{j=0}^{k-1} \mathrm{OPT}_f(\mathbb{1}_{S_j}) \\
&\leq k + (1+2\delta)\left(\delta\sum_{j=0}^{k-1}\mathrm{OPT}_f(\mathbb{1}_{S_j}\wedge \bar{y}^*) + \delta\sum_{j=0}^{k-1}\mathrm{OPT}_f\left(\mathbb{1}_{S_j}\wedge(\mathbb{1}_I - \bar{y}^*)\right)\right) \\
&\leq k + (1+2\delta)\left(\frac{4}{3}\cdot\bar{x}^0\cdot\mathbb{1}_{U_0\setminus\mathcal{C}_2} + 30\delta\mathrm{OPT} + \frac{4}{3}\cdot\bar{x}^0\cdot\mathbb{1}_{T_0\setminus\mathcal{C}_2} + \frac{1}{3}\cdot|\mathcal{M}| + 50\delta(\mathrm{OPT}+|\mathcal{M}|)\right) \\
&\leq k + (1+2\delta)\left(\frac{4}{3}\cdot\bar{x}^0\cdot\mathbb{1}_{\mathcal{C}\setminus\mathcal{C}_2} + \frac{1}{3}\cdot|\mathcal{M}| + 80\delta(\mathrm{OPT}+|\mathcal{M}|)\right) \\
&\leq \frac{4}{3}\cdot\bar{x}^0\mathbb{1}_{\mathcal{C}\setminus\mathcal{C}_2} + \frac{1}{3}\cdot|\mathcal{M}| + 90\delta(\mathrm{OPT}+|\mathcal{M}|)\ .
\end{aligned}$$

The first inequality uses (46), and the last inequality assumes $\mathrm{OPT} > \frac{k}{\delta}$. The number of configurations returned by the algorithm (assuming the statement of the lemmas hold) is

$$\begin{aligned}
|\mathcal{M}| + \sum_{j=1}^{k}\rho_j + \rho^* &\leq |\mathcal{M}| + \frac{4}{3}\cdot\bar{x}^0\cdot\mathbb{1}_{\mathcal{C}\setminus\mathcal{C}_2} + \frac{1}{3}\cdot|\mathcal{M}| + 90\delta(\mathrm{OPT}+|\mathcal{M}|) + 16\delta\mathrm{OPT} + 1 \\
&\leq \frac{4}{3}\cdot\bar{x}^0\cdot\mathbb{1}_{\mathcal{C}\setminus\mathcal{C}_2} + 110\delta\cdot\mathrm{OPT} + \left(\frac{4}{3}+90\delta\right)|\mathcal{M}| \\
&\leq \frac{4}{3}\cdot\bar{x}^0\cdot\mathbb{1}_{\mathcal{C}\setminus\mathcal{C}_2} + 110\delta\cdot\mathrm{OPT} + \left(\frac{4}{3}+90\delta\right)\cdot\left(\bar{x}^0\cdot\mathbb{1}_{\mathcal{C}_2} + \delta^2\mathrm{OPT}\right) \\
&\leq \left(\frac{4}{3}+90\delta\right)\|\bar{x}^0\| + 110\delta\mathrm{OPT} + 90\delta^3\mathrm{OPT} \\
&\leq \left(\frac{4}{3}+250\delta\right)\mathrm{OPT},
\end{aligned}$$

where the last inequality uses $\|\bar{x}\|^0 \leq (1+\delta^2)\mathrm{OPT}$. $\square$

4.1.3 Concentration

In this section we give the missing proofs of Section 4.1 and Section 4.1.1.

Proof of Lemma 4.8. Let $\mathcal{S} = \{\bar{u}^1, \ldots, \bar{u}^{\lfloor\varphi(\delta)\rfloor}\}$, where $\bar{u}^\ell$ is an $\mathcal{F}_j$-measurable random vector for $\ell \in [\varphi(\delta)]$ (in case $|\mathcal{S}| < \lfloor\varphi(\delta)\rfloor$ the same vector may appear several times in $\bar{u}^1, \ldots, \bar{u}^{\lfloor\varphi(\delta)\rfloor}$). As $\mathcal{S}$ is a $(\delta, \varphi(\delta))$ linear structure, it holds that

$$
\begin{aligned}
&\Pr\left(\forall r = j, \ldots, k: \quad \mathrm{OPT}_f\left(\bar{d} \wedge \mathbb{1}_{S_r}\right) \le (1-\delta)^{r-j}(1+10\delta)\|\bar{\lambda}\| + \varphi(\delta) + \delta^{10} \cdot \mathrm{OPT}\right) \\
&\quad \ge \Pr\left(\forall r = j, \ldots, k, \ell = 1, \ldots, \varphi(\delta): \quad (\mathbb{1}_{S_r} \wedge \bar{d}) \cdot \bar{u}^\ell \le (1-\delta)^{r-j} \cdot \bar{w} \cdot \bar{u}^\ell + \frac{\mathrm{OPT}}{\varphi^{10}(\delta)} \cdot \mathsf{tol}(\bar{u}^\ell)\right) \\
&\quad \ge \Pr\begin{pmatrix} \forall \ell = 1, \ldots, \varphi(\delta): & (\mathbb{1}_{S_j} \wedge \bar{d}) \cdot \bar{u}^\ell \le \bar{w} \cdot \bar{u}^\ell + \frac{1}{\varphi^{11}(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u}^\ell) \\ \forall \ell = 1, \ldots, \varphi(\delta),\ r = j, \ldots, k: & (\mathbb{1}_{S_r} \wedge \bar{d}) \cdot \bar{u}^\ell \le (1-\delta)^{r-j} \cdot (\mathbb{1}_{S_j} \wedge \bar{d}) \cdot \bar{u}^\ell + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)} \cdot \mathsf{tol}(\bar{u}^\ell) \end{pmatrix} \\
&\quad \ge \xi - \sum_{\ell=1}^{\lfloor\varphi(\delta)\rfloor} \Pr\left(\exists r \in \{j, \ldots, k\}: (\mathbb{1}_{S_r} \wedge \bar{d}) \cdot \bar{u}^\ell > (1-\delta)^{r-j} \cdot (\mathbb{1}_{S_j} \wedge \bar{d}) \cdot \bar{u}^\ell + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)} \cdot \mathsf{tol}(\bar{u}^\ell)\right) \\
&\quad \ge \xi - \varphi(\delta) \cdot \delta^{-2} \cdot \exp\left(-\frac{2 \cdot \delta^4 \cdot \left(\frac{\mathrm{OPT}^2}{\varphi^{11}(\delta)}\right)^2}{\mathrm{OPT}}\right) \\
&\quad \ge \xi - \varphi^2(\delta) \cdot \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \quad .
\end{aligned}
$$

The fourth equality follows from the union bound and the definition of ξ in (48). The fifth inequality is by Lemma 2.4. $\square$

The following technical lemma will be used to prove Lemma 4.9.

Lemma 4.18. *Let $j \in \{1, \ldots, k\}$ and $h \in \{2, \ldots, 2 \cdot \delta^{-1}\}$. Then*

$$
\Pr\left(\left|\mathbb{E}\left[\bar{x}^* \cdot \mathbb{1}_{T_{j,h}} \mid \mathcal{F}_{j-1}\right] - \bar{x}^* \cdot \mathbb{1}_{T_{j,h}}\right| > \delta^{20} \cdot \mathrm{OPT}\right) \le 2 \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT}\right) \quad .
$$

Proof. Let $\mathcal{V} \subseteq [0,1]^{\mathcal{C}}$ be the set of values that $\bar{x}^*$ can take, that is, $\mathcal{V} = \{\bar{x}^*(\omega) \mid \omega \in \Omega\}$. Since Ω is finite, it follows that $\mathcal{V}$ is finite as well. Furthermore, since $\sum_{C \in \mathcal{C}} \bar{x}^*_C \cdot C(i) \le \sum_{C \in \mathcal{C}} \bar{x}^0_C \cdot C(i) = 1$ for every $i \in I$, it follows that $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \le 1$ for every $\bar{x} \in \mathcal{V}$ and $i \in I$.

For any $U \subseteq \mathcal{C}$, $\rho = 1, \ldots, \mathrm{OPT}$ and $\bar{x} \in \mathcal{V}$ define $f_{U,\rho,\bar{x}} : \mathcal{C}^{\mathrm{OPT}} \to \mathbb{R}$ by

$$
f_{U,\rho,\bar{x}}\left(C_1, \ldots, C_{\mathrm{OPT}}\right) = \bar{x} \cdot \mathbb{1}_{\{C \in U \ \mid\ C \cap (\bigcup_{\ell=1}^{\rho} C_\ell) \cap L \neq \emptyset\}} = \sum_{C \in U} \bar{x}_C \cdot \mathbb{1}_{C \cap (\bigcup_{\ell=1}^{\rho} C_\ell) \cap L \neq \emptyset} \quad .
$$

Define $D = \{f_{U,\rho,\bar{x}} \mid U \subseteq \mathcal{C},\ \rho = 1, \ldots, \mathrm{OPT},\ \bar{x} \in \mathcal{V}\}$. It follows that D is a finite set.

Let $f_{U,\rho,\bar{x}} \in D$, $(C_1, \ldots, C_{\mathrm{OPT}})$, $(C'_1, \ldots, C'_{\mathrm{OPT}}) \in \mathcal{C}^{\mathrm{OPT}}$, and $r = 1, \ldots, \mathrm{OPT}$ such that $C_\ell = C'_\ell$ for $\ell = 1, \ldots, r-1, r+1, \ldots, \mathrm{OPT}$. If $r > \rho$ then $|f_{U,\rho,\bar{x}}(C_1, \ldots, C_{\mathrm{OPT}}) - f_{U,\rho,\bar{x}}(C'_1, \ldots, C'_{\mathrm{OPT}})| = 0$.

Otherwise, let $T = \bigcup_{\ell \in \{1,\ldots,\rho\}\setminus\{r\}} C_\ell = \bigcup_{\ell \in \{1,\ldots,\rho\}\setminus\{r\}} C'_\ell$. It holds that

$$\begin{aligned}
&\left| f_{U,\rho,\bar{x}}(C_1,\ldots,C_{\mathrm{OPT}}) - f_{U,\rho,\bar{x}}(C'_1,\ldots,C'_{\mathrm{OPT}}) \right| \\
&\qquad = \left| \bar{x} \cdot \left(\mathbb{1}_{\{C \in U \ \mid \ C \cap (T \cup C_r) \cap L \neq \emptyset\}} - \mathbb{1}_{\{C \in U \ \mid \ C \cap (T \cup C'_r) \cap L \neq \emptyset\}} \right) \right| \\
&\qquad = \left| \sum_{C \in U} \bar{x}_C \cdot \mathbb{1}_{C \cap (T \cup C'_r) \cap L = \emptyset} \cdot \mathbb{1}_{C \cap C_r \cap L \neq \emptyset} - \sum_{C \in U} \bar{x}_C \cdot \mathbb{1}_{C \cap (T \cup C_r) \cap L = \emptyset} \cdot \mathbb{1}_{C \cap C'_r \cap L \neq \emptyset} \right| \\
&\qquad \leq \max \left\{ \sum_{C \in U} \bar{x}_C \cdot \mathbb{1}_{C \cap (T \cup C'_r) \cap L = \emptyset} \cdot \mathbb{1}_{C \cap C_r \cap L \neq \emptyset}, \ \sum_{C \in U} \bar{x}_C \cdot \mathbb{1}_{C \cap (T \cup C_r) \cap L = \emptyset} \cdot \mathbb{1}_{C \cap C'_r \cap L \neq \emptyset} \right\} \\
&\qquad \leq \max \left\{ \sum_{C \in \mathcal{C}} \bar{x}_C \cdot \mathbb{1}_{C \cap C_r \cap L \neq \emptyset}, \ \sum_{C \in \mathcal{C}} \bar{x}_C \cdot \mathbb{1}_{C \cap C'_r \cap L \neq \emptyset} \right\} \quad .
\end{aligned}$$

Furthermore,

$$\sum_{C \in \mathcal{C}} \bar{x}_C \cdot \mathbb{1}_{C \cap C_r \cap L \neq \emptyset} \leq \sum_{C \in \mathcal{C}} \bar{x}_C \sum_{i \in C_r \cap L} C(i) = \sum_{i \in C_r \cap L} \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \leq |C_r \cap L| \leq 2 \cdot \delta^{-1},$$

and by a symmetric argument $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot \mathbb{1}_{C \cap C'_r \cap L \neq \emptyset} \leq 2 \cdot \delta^{-1}$. Thus,

$$\left| f_{U,\rho,\bar{x}}(C_1,\ldots,C_{\mathrm{OPT}}) - f_{U,\rho,\bar{x}}(C'_1,\ldots,C'_{\mathrm{OPT}}) \right| \leq 2 \cdot \delta^{-1} \quad .$$

That is, all functions in D are of $(2\delta^{-1})$-bounded difference.

Define $g = f_{U_{j-1,h},\rho_j,\bar{x}^*}$. Since $U_{j-1,h}$, ρ_j and $\bar{x}^*$ are $\mathcal{F}_{j-1}$-measurable, we have that g is a $\mathcal{F}_{j-1}$-measurable random function. For every $C \in \mathcal{C}$ it holds that $C \in T_{j,h}$ if and only if $C \in U_{j-1,h}$ and $C \cap L \cap \left(\bigcup_{\ell \in [\rho_j]} C^j_\ell \right) \neq \emptyset$. Thus,

$$g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) = \bar{x}^* \cdot \mathbb{1}_{\{C \in U_{j-1,h} \ \mid \ C \cap \left(\bigcup_{\ell=1}^{\rho_j} C^j_\ell\right) \cap L \neq \emptyset\}} = \bar{x}^* \cdot \mathbb{1}_{T_{j,h}} \quad .$$

Therefore,

$$\begin{aligned}
&\Pr\left(\left| \mathbb{E}\left[\bar{x}^* \cdot \mathbb{1}_{T_{j,h}} \ \middle| \ \mathcal{F}_{j-1} \right] - \bar{x}^* \cdot \mathbb{1}_{T_{j,h}} \right| > \delta^{20} \cdot \mathrm{OPT} \right) \\
&\quad = \Pr\left(\left| \mathbb{E}\left[g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) \ \middle| \ \mathcal{F}_{j-1} \right] - g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) \right| > \delta^{20} \cdot \mathrm{OPT} \right) \\
&\quad = \Pr\left(\mathbb{E}\left[g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) \ \middle| \ \mathcal{F}_{j-1} \right] - g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) > \delta^{20} \cdot \mathrm{OPT} \right) \\
&\qquad + \Pr\left(\mathbb{E}\left[-g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) \ \middle| \ \mathcal{F}_{j-1} \right] + g(C^j_1,\ldots,C^j_{\mathrm{OPT}}) > \delta^{20} \cdot \mathrm{OPT} \right) \\
&\quad \leq 2 \cdot \exp\left(-\frac{2 \cdot \delta^{40} \cdot \mathrm{OPT}^2}{(2\delta^{-1})^2 \cdot \mathrm{OPT}} \right) \leq 2 \cdot \exp\left(-\delta^{50} \cdot \mathrm{OPT} \right),
\end{aligned}$$

where the inequality follows from Lemma 2.3. ◻

The proof of Lemma 4.9 follows directly from Lemma 4.18.

Proof of Lemma 4.9. By the union bound, we have

$$\begin{aligned}
&\Pr\left(\forall j=1,\ldots,k, h=2,\ldots,2\cdot\delta^{-1}:\ \left|\mathbb{E}\left[\bar{x}^*\cdot\mathbb{1}_{T_{j,h}}\mid\mathcal{F}_{j-1}\right]-\bar{x}^*\cdot\mathbb{1}_{T_{j,h}}\right|\le\delta^{20}\cdot\mathrm{OPT}\right)\\
&\geq 1-\sum_{j=1}^{k}\sum_{h=2}^{2\cdot\delta^{-1}}\Pr\left(\left|\mathbb{E}\left[\bar{x}^*\cdot\mathbb{1}_{T_{j,h}}\mid\mathcal{F}_{j-1}\right]-\bar{x}^*\cdot\mathbb{1}_{T_{j,h}}\right|>\delta^{20}\cdot\mathrm{OPT}\right)\\
&\geq 1-k\cdot 2\cdot\delta^{-1}\cdot 2\cdot\exp\left(-\delta^{50}\cdot\mathrm{OPT}\right)\\
&\geq 1-\delta^{-10}\cdot\exp(-\delta^{50}\cdot\mathrm{OPT}),
\end{aligned}$$

where the second inequality follows from Lemma 4.18 and the last inequality uses $k\le\delta^{-2}$. $\square$

We use Lemma 4.9 to prove Lemma 4.10.

Proof of Lemma 4.10. For every $\varepsilon\in(0,0.1)$ and $h\in\mathbb{N}$, it holds that $\lim_{z\to\infty}\left(1-\frac{h}{z}\right)^{\lceil -z\cdot\ln(1-\varepsilon)\rceil}=(1-\varepsilon)^h$; thus, there is $M_{\varepsilon,h}>1$ such that for every $z>M_{\varepsilon,h}$ it holds that $\left(1-\frac{h}{z}\right)^{\lceil -z\cdot\ln(1-\varepsilon)\rceil}\geq(1-\varepsilon)^h-\varepsilon^{20}$. Define $\mu:(0,0.1)\to\mathbb{R}_+$ by $\mu(\varepsilon)=\max\left\{M_{\varepsilon,h}\mid h\in[2,2\cdot\varepsilon^{-1}]\cap\mathbb{N}\right\}$ for every $\varepsilon\in(0,0.1)$. Note that since the maximum is taken over a finite set of numbers, each greater than one, it follows that $\mu(\varepsilon)\in(1,\infty)$ for every $\varepsilon\in(0,0.1)$.

Assume the event in (49) occurs. Let $j\in\{1,\ldots,k\}$ and $h\in\{2,\ldots,2\delta^{-1}\}$. For any $C\in\mathcal{C}_h$ it holds that

$$\begin{aligned}
\Pr\left(C\in U_{j,h}\mid\mathcal{F}_{j-1}\right)&=\mathbb{1}_{C\in U_{j-1,h}}\cdot\Pr\left(\forall\ell\in 1,\ldots,\rho_j:\ C^j_\ell\cap C\cap L=\emptyset\ \middle|\ \mathcal{F}_{j-1}\right)\\
&=\mathbb{1}_{C\in U_{j-1,h}}\cdot\left(1-\frac{\sum_{C'\in\mathcal{C}}\bar{x}^j_{C'}\cdot\mathbb{1}_{C'\cap C\cap L\neq\emptyset}}{z_j}\right)^{\lceil -z_j\cdot\ln(1-\delta)\rceil}\\
&\geq\mathbb{1}_{C\in U_{j-1,h}}\cdot\left(1-\frac{h}{z_j}\right)^{\lceil -z_j\cdot\ln(1-\delta)\rceil}\\
&\geq\mathbb{1}_{\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}})>\mu(\delta)}\cdot\mathbb{1}_{C\in U_{j-1,h}}\cdot\left((1-\delta)^h-\delta^{20}\right)\ .
\end{aligned}\tag{65}$$

The first inequality holds, since, for every $C\in\mathcal{C}$,

$$\sum_{C'\in\mathcal{C}}\bar{x}^j_{C'}\cdot\mathbb{1}_{C'\cap C\cap L\neq\emptyset}\leq\sum_{C'\in\mathcal{C}}\bar{x}^j_{C'}\cdot\sum_{i\in C\cap L}C'(i)=\sum_{i\in C\cap L}\sum_{C'\in\mathcal{C}}\bar{x}^j_{C'}\cdot C(i)\leq h\ .$$

The last inequality in (65) holds by definition of μ and since $z_j\geq\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}})$.

We therefore have

$$\begin{aligned}
\mathbb{1}_{U_{j,h}}\cdot\bar{x}^*&=\mathbb{1}_{U_{j-1,h}}\cdot\bar{x}^*-\mathbb{1}_{T_{j,h}}\cdot\bar{x}^*\\
&\geq\mathbb{1}_{U_{j-1,h}}\cdot\bar{x}^*-\mathbb{E}\left[\mathbb{1}_{T_{j,h}}\cdot\bar{x}^*\mid\mathcal{F}_{j-1}\right]-\delta^{20}\cdot\mathrm{OPT}\\
&=\mathbb{E}\left[\mathbb{1}_{U_{j,h}}\cdot\bar{x}^*\mid\mathcal{F}_{j-1}\right]-\delta^{20}\cdot\mathrm{OPT}\\
&\geq\mathbb{1}_{\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}})>\mu(\delta)}\cdot\mathbb{1}_{U_{j-1,h}}\cdot\bar{x}^*\left((1-\delta)^h-\delta^{20}\right)-\delta^{20}\cdot\mathrm{OPT}\\
&\geq\mathbb{1}_{\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}})>\mu(\delta)}\cdot\mathbb{1}_{U_{j-1,h}}\cdot\bar{x}^*\cdot(1-\delta)^h-\delta^{19}\cdot\mathrm{OPT}\ .
\end{aligned}$$

The first inequality is due to (49), the second inequality follows from (65), and the last inequality uses $\mathbb{1}_{U_{j-1,h}}\cdot\bar{x}^*\leq\|\bar{x}^*\|\leq\|\bar{x}^0\|\leq 2\mathrm{OPT}$. Overall, we showed that

$$\mathbb{1}_{U_{j,h}}\cdot\bar{x}^*\geq\mathbb{1}_{\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}})>\mu(\delta)}\cdot\mathbb{1}_{U_{j-1,h}}\cdot\bar{x}^*\cdot(1-\delta)^h-\delta^{19}\cdot\mathrm{OPT}\tag{66}$$

for $j=1,\ldots,k$ and $h=2,\ldots,2\cdot\delta^{-1}$.

Claim 4.19. *For $h = 2, \dots, 2 \cdot \delta^{-1}$ and $j = 0, 1, \dots, k$ it holds that*

$$\bar{x}^* \cdot \mathbb{1}_{U_{j,h}} \geq (1-\delta)^{h \cdot j} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - j \cdot \delta^{19} \cdot \mathrm{OPT} \text{ or } \mathrm{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta) \ .$$

Proof. Fix $h \in \{2, \dots, 2 \cdot \delta^{-1}\}$. We show the claim by induction over j.
Base case: For $j = 0$ it clearly holds that $\bar{x}^* \cdot \mathbb{1}_{U_{0,h}} \geq (1-\delta)^{h \cdot 0} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - 0 \cdot \delta^{19} \cdot \mathrm{OPT}$.
Induction step: Assume the induction hypothesis holds for $j-1$. If $\mathrm{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta)$ then the statement holds for j. Otherwise, $\mathrm{OPT}_f(\mathbb{1}_{S_j}) > \mu(\delta)$, and so $\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}}) \geq \mathrm{OPT}_f(\mathbb{1}_{S_j}) > \mu(\delta)$. By the induction hypothesis, we have

$$\bar{x}^* \cdot \mathbb{1}_{U_{j-1,h}} \geq (1-\delta)^{h \cdot (j-1)} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - (j-1) \cdot \delta^{19} \cdot \mathrm{OPT} \ . \tag{67}$$

Therefore,

$$\begin{aligned}
\bar{x}^* \cdot \mathbb{1}_{U_{j,h}} &\geq \mathbb{1}_{\mathrm{OPT}_f(\mathbb{1}_{S_{j-1}}) > \mu(\delta)} \cdot \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* \cdot (1-\delta)^h - \delta^{19} \cdot \mathrm{OPT} \\
&= \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* \cdot (1-\delta)^h - \delta^{19} \cdot \mathrm{OPT} \\
&\geq (1-\delta)^h \left((1-\delta)^{h \cdot (j-1)} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - (j-1) \cdot \delta^{19} \cdot \mathrm{OPT} \right) - \delta^{19} \cdot \mathrm{OPT} \\
&\geq (1-\delta)^{h \cdot j} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - j \cdot \delta^{19} \cdot \mathrm{OPT} \ .
\end{aligned}$$

The first inequality holds by (66), and the second inequality is by (67). $\diamond$

By Claim 4.19, for $j = 1, \dots, k$ and $h = 2, \dots, 2 \cdot \delta^{-2}$, either $\mathrm{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta)$, or

$$\bar{x}^* \cdot \mathbb{1}_{U_{j,h}} \geq (1-\delta)^{h \cdot j} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - j \cdot \delta^{19} \cdot \mathrm{OPT} \geq (1-\delta)^{h \cdot j} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - \delta^{10} \cdot \mathrm{OPT},$$

as required (the last inequality uses $j \leq k \leq \delta^{-2}$). Since we assumed (49) occurs, this property holds with probability at least $1 - \delta^{-10} \cdot \exp(-\delta^{50} \cdot \mathrm{OPT})$ by Lemma 4.9. $\square$

We now proceed to the proof of Claim 4.14. We use the same notation as in the proof of Lemma 4.11, where the claim is stated.

Proof of Claim 4.14. As in the proof of Lemma 4.18, let $\mathcal{V} \subseteq [0,1]^{\mathcal{C}}$ be all the values $\bar{x}^*$ can take (formally, $\mathcal{V} = \{\bar{x}^*(\omega) \mid \omega \in \Omega\}$). It follows that $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \leq 1$ for every $i \in I$ and $\bar{x} \in \mathcal{V}$. Also, let $A \subseteq \mathbb{R}^I_{\geq 0}$ be the set of all values the vectors in $\mathcal{S}_j$ can take (formally, $A = \{\bar{u} \mid \exists \omega \in \Omega : \ \bar{u} \in \mathcal{S}_j(\omega)\}$) As Ω is finite, it follows that $\mathcal{V}$ and A are finite.

For any $U \subseteq \mathcal{C}$, $S \subseteq I$, $\bar{x} \in \mathcal{V}$, $\rho \in [\mathrm{OPT}]$ and $\bar{u} \in A$, we define $f_{U,S,\bar{x},\rho,\bar{u}} : \mathcal{C}^{\mathrm{OPT}} \to \mathbb{R}$ by

$$f_{U,S,\bar{x},\rho,\bar{u}}(C_1, \dots, C_{\mathrm{OPT}}) = \begin{cases} \dfrac{1}{\mathsf{tol}(\bar{u})} \cdot \displaystyle\sum_{C \in U} \bar{x}_C \cdot \mathbb{1}_{C \cap (\bigcup_{\ell \in [\rho]} C_\ell) \cap L \neq \emptyset} \cdot \sum_{i \in C \cap S} \mathbb{1}_{i \notin \bigcup_{\ell = 1, \dots, \rho} C_\ell} \cdot \bar{u}_i & \mathsf{tol}(\bar{u}) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Let $D = \{f_{U,S,\bar{x},\rho,\bar{u}} \mid U \subseteq \mathcal{C}, \ S \subseteq I, \ \bar{x} \subseteq \mathcal{V}, \ \rho \in [\mathrm{OPT}], \ \bar{u} \in A\}$. It follows that D is finite.

Let $f_{U,S,\bar{x},\rho,\bar{u}} \in D$, $(C_1, \dots, C_{\mathrm{OPT}})$, $(C'_1, \dots, C'_{\mathrm{OPT}}) \in \mathcal{C}^{\mathrm{OPT}}$ and $r \in \{1, \dots, \mathrm{OPT}\}$ be such that $C_\ell = C'_\ell$ for $\ell = 1, \dots, r-1, r+1, \dots \mathrm{OPT}$. If $\mathsf{tol}(\bar{u}) = 0$ or $r > \rho$,

$$\left| f_{U,S,\bar{x},\rho,\bar{u}}(C_1, \dots, C_{\mathrm{OPT}}) - f_{U,S,\bar{x},\rho,\bar{u}}(C'_1, \dots, C'_{\mathrm{OPT}}) \right| = 0 \ .$$

Otherwise, let $T = \bigcup_{\ell\in\{1,\ldots,\rho\}\setminus\{r\}} C_\ell = \bigcup_{\ell\in\{1,\ldots,\rho\}\setminus\{r\}} C'_\ell$. Then

$$
\begin{aligned}
&\left| f_{U,S,\bar{x},\rho,\bar{u}}(C_1,\ldots,C_{\mathrm{OPT}}) - f_{U,S,\bar{x},\rho,\bar{u}}(C'_1,\ldots,C'_{\mathrm{OPT}}) \right| \\
&= \frac{1}{\mathsf{tol}(\bar{u})} \cdot \left| \sum_{C\in U} \bar{x}_C \cdot \mathbb{1}_{C\cap(T\cup C_r)\cap L\neq\emptyset} \cdot \sum_{i\in C\cap S} \mathbb{1}_{i\notin T\cup C_r} \cdot \bar{u}_i - \sum_{C\in U} \bar{x}_C \cdot \mathbb{1}_{C\cap(T\cup C'_r)\cap L\neq\emptyset} \cdot \sum_{i\in C\cap S} \mathbb{1}_{i\notin T\cup C'_r} \cdot \bar{u}_i \right| \\
&= \frac{1}{\mathsf{tol}(\bar{u})} \cdot \left| \sum_{C\in U} \sum_{i\in C\cap S} \bar{x}_C \cdot \bar{u}_i \cdot \left(\mathbb{1}_{C\cap(T\cup C_r)\cap L\neq\emptyset} \cdot \mathbb{1}_{i\notin T\cup C_r} - \mathbb{1}_{C\cap(T\cup C'_r)\cap L\neq\emptyset} \cdot \mathbb{1}_{i\notin T\cup C'_r} \right) \right| \\
&\le \frac{1}{\mathsf{tol}(\bar{u})} \cdot \sum_{C\in U} \sum_{i\in C\cap S} \bar{x}_C \cdot \bar{u}_i \cdot \left| \mathbb{1}_{C\cap(T\cup C_r)\cap L\neq\emptyset} \cdot \mathbb{1}_{i\notin T\cup C_r} - \mathbb{1}_{C\cap(T\cup C'_r)\cap L\neq\emptyset} \cdot \mathbb{1}_{i\notin T\cup C'_r} \right| \\
&\le \frac{1}{\mathsf{tol}(\bar{u})} \cdot \sum_{C\in U} \sum_{i\in C\cap S} \bar{x}_C \cdot \bar{u}_i \cdot \left(\mathbb{1}_{C\cap(C'_r\cup C_r)\cap L\neq\emptyset} + \mathbb{1}_{i\in C_r\cup C'_r} \right) \\
&\le \frac{1}{\mathsf{tol}(\bar{u})} \sum_{C\in U} \mathbb{1}_{C\cap(C'_r\cup C_r)\cap L\neq\emptyset} \cdot \bar{x}_C \cdot \sum_{i\in C} \bar{u}_i \;+\; \frac{1}{\mathsf{tol}(\bar{u})} \cdot \sum_{i\in C_r\cup C'_r} \bar{u}_i \cdot \sum_{C\in U} \bar{x}_C \cdot C(i) \\
&\le \frac{1}{\mathsf{tol}(\bar{u})} \cdot \mathsf{tol}(\bar{u}) \cdot 4 \cdot \delta^{-1} + \frac{1}{\mathsf{tol}(\bar{u})} \sum_{i\in C_r\cup C'_r} \bar{u}_i \\
&\le 4 \cdot \dot{\delta}^{-1} + \frac{1}{\mathsf{tol}(\bar{u})} \cdot 2 \cdot \mathsf{tol}(\bar{u}) \\
&\le \delta^{-2},
\end{aligned}
$$

where the fourth inequality uses

$$
\sum_{C\in U} \mathbb{1}_{C\cap(C'_r\cup C_r)\cap L\neq\emptyset} \cdot \bar{x}_C \le \sum_{i\in(C_r\cup C'_r)\cap L} \sum_{C\in\mathcal{C}} \bar{x}_C \cdot C(i) \le \sum_{i\in(C_r\cup C'_r)\cap L} 1 \le 4\cdot\delta^{-1} \;.
$$

We conclude that all functions in D are of δ^{-2}-bounded difference.

Recall $\mathcal{S}_j$ is a $(\delta,\varphi(\delta))$-linear structure of $\bar{\lambda}^j$. Since $\bar{\lambda}^j$ is $\mathcal{F}_{j-1}$-measurable, it follows that $\mathcal{S}_j$ is also $\mathcal{F}_{j-1}$-measurable. As in the proof of Lemma 4.8, we denote $\mathcal{S}_j = \{\bar{u}^1,\ldots,\bar{u}^{\lfloor\varphi(\delta)\rfloor}\}$ where $\bar{u}^s$ is an $\mathcal{F}_{j-1}$-measurable random vector for $s=1,\ldots,\lfloor\varphi(\delta)\rfloor$ (in case $|\mathcal{S}_j| < \lfloor\varphi(\delta)\rfloor$ the same vector may appear several times in $\bar{u}^1,\ldots,\bar{u}^{\lfloor\varphi(\delta)\rfloor}$).

For $s=1,\ldots,\lfloor\varphi(\delta)\rfloor$ define a random function $g^s = f_{U_{j-1},S_{j-1},\bar{x}^*,\rho_j,\bar{u}^s}$. Since $U_{j-1}, S_{j-1}, \bar{x}^*, \rho_j$ and $\bar{u}^s$ are all $\mathcal{F}_{j-1}$-measurable, it follows that g^s is $\mathcal{F}_{j-1}$-measurable as well. Furthermore,

$$
\begin{aligned}
\mathsf{tol}(\bar{u}^s)\cdot g^s(C^j_1,\ldots,C^j_{\mathrm{OPT}}) &= \sum_{C\in U_{j-1}} \bar{x}^*_C \cdot \mathbb{1}_{C\cap\left(\bigcup_{\ell=1,\ldots,\rho_j} C^j_\ell\right)\cap L\neq\emptyset} \cdot \sum_{i\in C\cap S_{j-1}} \mathbb{1}_{i\notin\bigcup_{\ell\in[\rho_j]} C^j_\ell} \cdot \bar{u}^s_i \\
&= \sum_{i\in I} \mathbb{1}_{i\in S_j} \cdot \bar{u}^s_i \cdot \sum_{C\in T_j} \bar{x}^*_C \cdot C(i) = \sum_{i\in I} \mathbb{1}_{i\in S_j} \cdot \bar{u}^s_i \cdot \bar{d}^j_i = (\mathbb{1}_{S_j} \wedge \bar{d}^j)\cdot\bar{u}^s,
\end{aligned}
$$

where the third equality follows from the definition of $\bar{d}^j$. Thus, for $s=1,\ldots,\lfloor\varphi(\delta)\rfloor$ it holds that

$$
\begin{aligned}
&\Pr\left((\mathbb{1}_{S_j}\wedge\bar{d}^j)\cdot\bar{u}^s \;>\; \mathbb{E}\left[\bar{u}^s\cdot\left(\bar{d}^j\wedge\mathbb{1}_{S_j}\right) \;\middle|\; \mathcal{F}_{j-1}\right] + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\cdot\mathsf{tol}(\bar{u}) \right) \\
&= \Pr\left(g^s(C^j_1,\ldots,C^j_{\mathrm{OPT}}) \;>\; \mathbb{E}\left[g^s(C^j_1,\ldots,C^j_{\mathrm{OPT}}) \;\middle|\; \mathcal{F}_{j-1}\right] + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)} \right) \\
&\le \exp\left(-\frac{2\cdot\left(\frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\right)^2}{\delta^{-4}\cdot\mathrm{OPT}} \right) \le \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)} \right),
\end{aligned}
$$

where the last inequality is by Lemma 2.3.

Thus, using the union bound we have that

$$\begin{aligned}
&\Pr\left(\forall \bar{u}\in\mathcal{S}_j:\ (\mathbb{1}_{S_j}\wedge \bar{d}^j)\cdot\bar{u}\ \le\ \mathbb{E}\left[\bar{u}\cdot(\bar{d}^j\wedge\mathbb{1}_{S_j})\ \middle|\ \mathcal{F}_{j-1}\right]+\frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\cdot\mathsf{tol}(\bar{u})\right)\\
&\ge 1-\sum_{s=1}^{\lfloor\varphi(\delta)\rfloor}\Pr\left((\mathbb{1}_{S_j}\wedge\bar{d}^j)\cdot\bar{u}^s\ >\ \mathbb{E}\left[\bar{u}^s\cdot(\bar{d}^j\wedge\mathbb{1}_{S_j})\ \middle|\ \mathcal{F}_{j-1}\right]+\frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\cdot\mathsf{tol}(\bar{u})\right)\\
&\ge 1-\varphi(\delta)\cdot\exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right)\ .
\end{aligned}$$

◇

It remains to prove Lemma 4.7 and Claim 4.17. We use $G=(L,E)$ to denote the δ-matching graph of (I,v), and $P_{\mathcal{M}}(G)$ to denote the matching polytope of G. Both proofs rely on the concentration bounds of SampleMatching given below.

Lemma 4.20 ([CVZ11])**.** *Let $\bar{\beta}\in P_{\mathcal{M}}(G)$ and $\gamma>0$. Also, denote $\mathcal{M}=\mathsf{SampleMatching}(\bar{\beta},\gamma)$. Then $\mathcal{M}$ is a matching, and for any $\bar{a}\in[0,1]^E$ the following holds:*

1. $\Pr(e\in\mathcal{M})=(1-\gamma)\bar{\beta}_e$ *for any* $e\in E$.
2. *For any* $\xi\le\mathbb{E}\left[\sum_{e\in\mathcal{M}}\bar{a}_e\right]$ *and* $\varepsilon>0$, *it holds that* $\Pr\left(\sum_{e\in\mathcal{M}}\bar{a}_e\le(1-\varepsilon)\cdot\xi\right)\le\exp\left(-\frac{\xi\cdot\varepsilon^2\cdot\gamma}{20}\right)$.
3. *For any* $\xi\ge\mathbb{E}\left[\sum_{e\in\mathcal{M}}\bar{a}_e\right]$ *and* $\varepsilon>0$, *it holds that* $\Pr\left(\sum_{e\in\mathcal{M}}\bar{a}_e\ge(1+\varepsilon)\cdot\xi\right)\le\exp\left(-\frac{\xi\cdot\varepsilon^2\cdot\gamma}{20}\right)$.

Proof of Lemma 4.7. As $\mathcal{M}=\mathsf{SampleMatching}(\mathcal{E}(\bar{x}^0),\delta^4)$, it follows that

$$\mathbb{E}\left[|\mathcal{M}|\right]=\sum_{e\in E}\Pr(e\in\mathcal{M})=(1-\delta^4)\cdot\sum_{e\in E}\mathcal{E}_e(\bar{x}^0)=(1-\delta^4)\cdot\sum_{e\in E}\ \sum_{C\in\mathcal{C}\text{ s.t. }e\in C}\bar{x}^0_C=(1-\delta^4)\cdot\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0\ .$$

If $\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0=0$, then $|\mathcal{M}|=0$, and the statement of the lemma holds.

Otherwise, by Lemma 4.20,

$$\begin{aligned}
\Pr\left(|\mathcal{M}|>\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0+\delta^2\cdot\mathrm{OPT}\right)&=\Pr\left(|\mathcal{M}|>\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0\cdot\left(1+\frac{\delta^2\cdot\mathrm{OPT}}{\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0}\right)\right)\\
&\le\exp\left(-\frac{1}{20}\cdot\delta^4\cdot(\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0)\cdot\left(\frac{\delta^2\cdot\mathrm{OPT}}{\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0}\right)^2\right)\le\exp\left(-\delta^{10}\cdot\mathrm{OPT}\right),
\end{aligned}$$

where the last inequality uses $\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0\le(1+\delta^2)\mathrm{OPT}\le 2\mathrm{OPT}$. Therefore,

$$\Pr\left(|\mathcal{M}|\le\mathbb{1}_{\mathcal{C}_2}\cdot\bar{x}^0+\delta^2\cdot\mathrm{OPT}\right)\ge 1-\exp\left(-\delta^{10}\cdot\mathrm{OPT}\right)\ .$$

□

Proof of Claim 4.17. We use the same notation as in the proof of Lemma 4.16, where the claim is stated. If $\mathsf{tol}(\bar{u})=0$ the claim trivially holds. Thus, we may assume that $\mathsf{tol}(\bar{u})\neq\emptyset$.

Observe that

$$\bar{w}\cdot\bar{u}-\bar{d}\cdot\bar{u}=\sum_{i\in I}\left(\bar{w}_i-\bar{d}_i\right)\bar{u}_i=\sum_{i\in L}\left(\mathbb{1}_{i\notin S_0}-(1-\delta^4)\cdot\bar{y}^{\mathcal{M}}_i\right)\bar{u}_i=\sum_{i\in L}\mathbb{1}_{i\notin S_0}\cdot\bar{u}_i-\mathbb{E}\left[\sum_{i\in L}\mathbb{1}_{i\notin S_0}\cdot\bar{u}_i\right],$$

where the second equality is by (59) and (60), and the last equality is by Lemma 4.15. Furthermore,

$$\sum_{i\in L} \mathbb{1}_{i\notin S_0} \cdot \bar{u}_i = \sum_{\{i_1,i_2\}\in\mathcal{M}} (\bar{u}_{i_1} + \bar{u}_{i_2}) \quad .$$

Thus,

$$\begin{aligned}
&\Pr\left(\bar{d}\cdot\bar{u} > \bar{w}\cdot\bar{u} + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\mathsf{tol}(\bar{u})\right)\\
&= \Pr\left(\sum_{i\in L}\mathbb{1}_{i\notin S_0}\cdot\bar{u}_i < \mathbb{E}\left[\sum_{i\in L}\mathbb{1}_{i\notin S_0}\cdot\bar{u}_i\right] - \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\cdot\mathsf{tol}(\bar{u})\right)\\
&= \Pr\left(\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})} < \mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right] - \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\right) \qquad (68)\\
&\le \exp\left(-\frac{1}{20}\cdot\delta^4\cdot\mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right]\cdot\left(\frac{\mathrm{OPT}}{\varphi^{11}(\delta)\cdot\mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right]}\right)^2\right)\\
&\le \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \quad .
\end{aligned}$$

The first inequality is by Lemma 4.20; observe that $\mathcal{M}\subseteq E\subseteq\mathcal{C}$, therefore $\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\le 1$ for any $\{i_1,i_2\}\in E$. The last inequality uses

$$\mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right] \le \frac{|L|}{2} \le \delta^{-1}\cdot\mathrm{OPT} \quad .$$

We implicitly assumed in (68) that $\mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right]\neq 0$. In case $\mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right]=0$, we have $\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}=0$, and

$$\begin{aligned}
\Pr\left(\bar{d}\cdot\bar{u} > \bar{w}\cdot\bar{u} + \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\mathsf{tol}(\bar{u})\right) &= \Pr\left(\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})} < \mathbb{E}\left[\sum_{\{i_1,i_2\}\in\mathcal{M}}\frac{\bar{u}_{i_1}+\bar{u}_{i_2}}{\mathsf{tol}(\bar{u})}\right] - \frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\right)\\
&= \Pr\left(0 < -\frac{\mathrm{OPT}}{\varphi^{11}(\delta)}\right) = 0 \le \exp\left(-\frac{\mathrm{OPT}}{\varphi^{25}(\delta)}\right) \quad .
\end{aligned}$$

$\diamond$

4.2 Proof of the Structural Lemma

In this section we give the proof of Lemma 4.2. Let $\delta\in(0,0.1)$ such that $\delta^{-1}\in\mathbb{N}$, and let (I,v) be a δ-2VBP instance. As in Section 4.1, we use $\mathrm{OPT}=\mathrm{OPT}(I,v)$.

We first need to construct the set $\mathcal{S}^*\subseteq\mathbb{R}^I_{\ge 0}$. The construction is technical; its components will become clearer below. The terms $\preceq_d$, $I_{d,j}$, h and $\hat{d}$ defined as part of the construction of $\mathcal{S}^*$ are also used in the construction of the linear structure $\mathcal{S}$.

Let $\succeq^*$ be an arbitrary total order[12] over I. For $d\in\{1,2\}$ we define a total order $\succeq_d$ on I by $i_1\succeq_d i_2$ if and only if $v_d(i_1)>v_d(i_2)$ or ($v_d(i_1)=v_d(i_2)$ and $i_1\succeq^* i_2$). Let $h=\delta^{-2}$. For

[12] We refer the reader to Appendix B.2 of Cormen et al. [CLRS01] for a formal definition of total order.

any $d \in \{1,2\}$ and $j = 1, \ldots, 2h$ we define a set $I_{d,j} = \left\{ i \in L \mid \frac{\delta^2}{2} \cdot (j-1) < v_d(i) \leq \frac{\delta^2}{2} \cdot j \right\}$. The construction of the linear structure $\mathcal{S}$ implicitly rounds the volume in dimension d of items in $I_{d,j}$ to $j \cdot \frac{\delta^2}{2}$, and applies *fractional grouping* to round the volume of the items in the dimension other than d, i.e., $\hat{d} = 3 - d$. For $d \in \{1,2\}$ define $\mathcal{S}_d^* = \left\{ \mathbb{1}_{\{i \in I_{d,j} \mid q_1 \preceq_{\hat{d}} i \preceq_{\hat{d}} q_2\}} \;\middle|\; j \in [2h],\ q_1, q_2 \in L \right\}$. The set $\mathcal{S}_d^*$ contains an indicator vector for every possible group which may be generated by the fractional grouping for $I_{d,j}$. Finally, the set $\mathcal{S}^*$ is defined by $\mathcal{S}^* = \left\{ \bar{u}^1 \wedge \bar{u}^2 \;\middle|\; \bar{u}^1 \in \mathcal{S}_1^*,\ \bar{u}^2 \in \mathcal{S}_2^* \right\}$. Observe that $|\mathcal{S}^*| \leq |\mathcal{S}_1^*| \cdot |\mathcal{S}_2^*| \leq \left(2h \cdot |L|^2\right)^2 = \delta^{-5} \cdot |L|^4 \leq \varphi(\delta) \cdot |L|^4$.

Let $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ be a small-items integral vector with δ-slack, and let $\bar{w} \in [0,1]^I$ be the coverage of $\bar{\lambda}$. In Section 4.2.1 we construct the linear structure $\mathcal{S}$ of $\bar{\lambda}$, and in Section 4.2.2 we show the structure indeed satisfies the requirements in Definition 4.1. The construction and proof of correctness rely on a technical *refinement* lemma whose proof is given in Section 4.2.3.

4.2.1 Construction of $\mathcal{S}$

Our construction uses a partition of $\bar{\lambda}$ into two parts: $\bar{\lambda}^1$ and $\bar{\lambda}^2$, such that for any $d \in \{1,2\}$ and $C \in \operatorname{supp}(\bar{\lambda}^d)$ it holds that C has δ-slack in dimension d. Formally, we define $\bar{\lambda}^1 \in [0,1]^{\mathcal{C}^*}$ by

$$\forall C \in \mathcal{C}^* : \qquad \bar{\lambda}^1_C = \begin{cases} \bar{\lambda}_C & \text{if } C \text{ has } \delta\text{-slack in dimension } 1 \\ 0 & \text{otherwise} \end{cases}$$

Also, we define $\bar{\lambda}^2 \in [0,1]^{\mathcal{C}^*}$ by $\bar{\lambda}^2 = \bar{\lambda} - \bar{\lambda}^1$. Indeed, as $\bar{\lambda}$ is with δ-slack, for every $d \in \{1,2\}$ and $C \in \operatorname{supp}(\bar{\lambda}^d)$, it holds that C has δ-slack in dimension d. For $d \in \{1,2\}$ let $\bar{w}^d$ be the coverage of $\bar{\lambda}^d$.

As mentioned above, for each $d \in \{1,2\}$ we implicitly give a rounding scheme for the large items, in which the volume in dimension d of all items in $i \in I_{d,j}$ is rounded up to $j \cdot \frac{\delta^2}{2}$. The slack of configurations in $\operatorname{supp}(\bar{\lambda})$ is used to compensate for the possible volume increase. In the other dimension, $\hat{d}$, we apply *fractional grouping*, defined as follows.

Definition 4.21. *Let $E \neq \emptyset$ be a finite set, $\bar{\gamma} \in [0,1]^E$, $\succeq$ be a total order over E and $\xi \in \mathbb{N}_+$. A partition $G_1, \ldots, G_\tau$ of E is a ξ-*fractional grouping* with respect to $\bar{\gamma}$ and $\succeq$ if the following conditions hold:*

1. *For every $1 \leq \ell_1 < \ell_2 \leq \tau$, $i_1 \in G_{\ell_1}$ and $i_2 \in G_{\ell_2}$ it holds that $i_1 \succeq i_2$.*
2. *For $\ell = 1, \ldots, \tau - 1$ it holds that $\mathbb{1}_{G_\ell} \cdot \bar{\gamma} \geq \frac{\|\bar{\gamma}\|}{\xi}$.*
3. *For $\ell = 1, \ldots, \tau$ it holds that $\mathbb{1}_{G_\ell} \cdot \bar{\gamma} \leq \frac{\|\bar{\gamma}\|}{\xi} + 1$.*

The proof of the next lemma utilizes arguments from Fairstein et al. [FKS21].

Lemma 4.22. *For any finite set $E \neq \emptyset$, $\bar{\gamma} \in [0,1]^E$, a total order $\succeq$ over E and $\xi \in \mathbb{N}_+$, there is a ξ-fractional grouping $G_1, \ldots, G_\tau$ of E with respect to $\bar{\gamma}$ and $\succeq$ for which $\tau \leq \xi$.*

Proof. If $\bar{\gamma} = \mathbf{0}$ then the partition $G_1 = E$ is a ξ-fractional grouping. We henceforth assume $\bar{\gamma} \neq \mathbf{0}$.

Assume, without loss of generality, that $E = \{1, 2, \ldots, \nu\} = [\nu]$ and $a \succeq b$ if and only if $a \leq b$. Define a sequence $(q_\ell)_{\ell=0}^{\infty}$ by $q_0 = 0$, and $q_\ell = \min \left\{ e \in E \;\middle|\; \sum_{f=q_{\ell-1}+1}^{e} \bar{\gamma}_f > \frac{\|\bar{\gamma}\|}{\xi} \right\} \cup \{\nu\}$. Also, define $\tau = \min\{\ell \in \mathbb{N} \mid q_\ell = \nu\}$. Since $\|\bar{\gamma}\| > 0$, it follows that $(q_\ell)_{\ell=0}^{\tau}$ is monotonically increasing.

We define $G_\ell = \{e \in E \mid q_{\ell-1} < e \leq q_\ell\}$. Then $G_\ell = \{1, \ldots, q_\ell\} \setminus \{1, \ldots, q_{\ell-1}\}$ for $\ell = 1, \ldots, \tau$. As $q_0 = 0$, $q_\tau = \nu$ and $(q_\ell)_{\ell=0}^{\tau}$ is monotonically increasing, it follows that $G_1, \ldots, G_\tau$ is a partition of E. Clearly, for any $1 \le \ell_1 < \ell_2 \le \tau$, $i_1 \in G_{\ell_1}$ and $i_2 \in G_{\ell_2}$ it holds that $i_1 \le q_{\ell_1} \le q_{\ell_2 - 1} < i_2$ thus $i_1 \succeq i_2$.

Let $\ell \in \{1, \ldots, \tau\}$. By definition of q_ℓ it holds that $\sum_{f=q_{\ell-1}+1}^{q_\ell - 1} \bar{\gamma}_f \le \frac{\|\bar{\gamma}\|}{\xi}$. Hence, as $\bar{\gamma}_{q_\ell} \le 1$, it also holds that $\mathbb{1}_{G_\ell} \cdot \bar{\gamma} = \sum_{e \in G_\ell} \bar{\gamma}_e = \bar{\gamma}_{q_\ell} + \sum_{e=q_{\ell-1}+1}^{q_\ell - 1} \bar{\gamma}_e \le \frac{\|\bar{\gamma}\|}{\xi} + 1$.

Let $\ell \in \{1, \ldots, \tau - 1\}$. Then $q_\ell \neq \nu$ and $q_\ell = \min \left\{ e \in E \;\middle|\; \sum_{f=q_{\ell-1}+1}^{e} \bar{\gamma}_f > \frac{\|\bar{\gamma}\|}{\xi} \right\}$. Therefore, $\mathbb{1}_{G_\ell} \cdot \bar{\gamma} = \sum_{e \in G_\ell} \bar{\gamma}_e = \sum_{e=q_{\ell-1}+1}^{q_\ell} \bar{\gamma}_e > \frac{\|\bar{\gamma}\|}{\xi}$.

Thus, we showed that $G_1, \ldots, G_\tau$ is a ξ-fractional grouping of E with respect to $\bar{\gamma}$ and $\succeq$. It also holds that

$$\|\bar{\gamma}\| = \sum_{e \in E} \bar{\gamma}_e = \sum_{\ell=1}^{\tau} \sum_{e \in G_\ell} \bar{\gamma}_e \ge \sum_{\ell=1}^{\tau-1} \sum_{e \in G_\ell} \bar{\gamma}_e > \sum_{\ell=1}^{\tau-1} \frac{\|\bar{\gamma}\|}{\xi} = (\tau - 1) \frac{\|\bar{\gamma}\|}{\xi} \;.$$

Hence, $\tau - 1 < \xi$, and as both τ and ξ are integral it follows that $\tau \le \xi$. This completes the proof. $\square$

For any $d \in \{1, 2\}$ and $j = 1, \ldots, 2h$ define a vector $\bar{\gamma}^{d,j} \in [0,1]^{I_{d,j}}$ by $\bar{\gamma}^{d,j}_i = \bar{w}^d_i$ for $i \in I_{d,j}$. By Lemma 4.22, for any $d \in \{1, 2\}$ and $j = 1, \ldots, 2h$ such that $I_{d,j} \neq \emptyset$ there is an h-fractional grouping $\left(G^{d,j}_\ell\right)_{\ell=1}^{\tau_{d,j}}$ of $I_{d,j}$ with respect to $\bar{\gamma}^{d,j}$ and the total order $\succeq_{\hat{d}}$ with $\tau_{d,j} \le h$. For $d \in \{1, 2\}$ let $\mathcal{G}_d = \{(j, \ell) \mid j \in [2h],\; I_{d,j} \neq \emptyset \text{ and } \ell \in [\tau_{d,j}]\}$. It follows that $\mathcal{G}_1, \mathcal{G}_2 \subseteq \{1, \ldots, 2h\} \times \{1, \ldots, h\}$ and thus $|\mathcal{G}_1|,\; |\mathcal{G}_2| \le 2\delta^{-4}$.

Our objective is to add to the structure $\mathcal{S}$ vectors $\bar{u}$ to ensure that if $\bar{z} \in [0,1]^I$ satisfies (45) then we can decompose $\bar{z} \wedge \mathbb{1}_L$ to $\bar{z}^1, \bar{z}^2 \in [0,1]^I$ such that $\bar{z} \wedge \mathbb{1}_L = \bar{z}^1 + \bar{z}^2$ and $\bar{z}^d \cdot \mathbb{1}_{G^{d,j}_\ell} \lesssim \beta \cdot \bar{w}^d \cdot \mathbb{1}_{G^{d,j}_\ell}$ for any $d \in \{1, 2\}$ and $(j, \ell) \in \mathcal{G}_d$. This can be intuitively interpreted as a decrease in demand for items in $G^{d,j}_\ell$ by a factor of β. As we have a rounding scheme for each dimension, an item $i \in L$ may belong to two groups $G^{d,j}_\ell$- one from the scheme for dimension 1 and another from the scheme of dimension 2. We therefore add into $\mathcal{S}$ vectors which represent the intersection of each pair of such groups, and therefore impose a decrease in demand by a factor of β for each intersection.

Formally, our linear structure will contain the set $\mathcal{S}_{\text{large}}$, which we define as

$$\mathcal{S}_{\text{large}} = \left\{ \mathbb{1}_{G^{1,j_1}_{\ell_1}} \wedge \mathbb{1}_{G^{2,j_2}_{\ell_2}} \;\middle|\; (j_1, \ell_1) \in \mathcal{G}_1,\; (j_2, \ell_2) \in \mathcal{G}_2 \right\} \;. \tag{69}$$

In Section 4.2.2 we show that if $S_{\text{large}} \subseteq \mathcal{S}$ and $\bar{z}$ satisfies (45) then we can find the decomposition $\bar{z}^1$ and $\bar{z}^2$ as mentioned above. Furthermore, to show the correctness of the structure we (implicitly) use a *shifting* argument (see, e.g., [FL81]) in which items in $G^{d,j}_\ell$ take the place of items in $G^{d,j}_{\ell-1}$.

We use the rounding schemes for the large items to define a *type* for each configuration. We then fractionally associate each small item $i \in I \setminus L$ with the various types, and use this association as a basis for the linear structure. For $d \in \{1, 2\}$, the *d-type* of a multi-configuration $C \in \mathcal{C}^*$, denoted by $\mathsf{T}^d(C)$, is the vector $\bar{t} \in \mathbb{N}^{\mathcal{G}_d}$ defined by $\bar{t}_{(j,\ell)} = \sum_{i \in G^{d,j}_\ell} C(i)$ for any $(j, \ell) \in \mathcal{G}_d$. That is, $\bar{t}_{(j,\ell)}$ is the number of items in C which belong to $G^{d,j}_\ell$. Since the set $G^{d,j}_\ell$ contains only large items, it follows that $\bar{t}_{(j,\ell)} \le 2\delta^{-1}$. Let $\mathcal{T}_d = \left\{ \mathsf{T}^d(C) \mid C \in \mathcal{C}^* \right\}$ be the set of all possible d-types. It follows that $\mathcal{T}_d \subseteq \{0, 1, \ldots, 2 \cdot \delta^{-1}\}^{\mathcal{G}_d}$, and therefore $|\mathcal{T}_d| \le \left(1 + 2 \cdot \delta^{-1}\right)^{2 \cdot \delta^{-4}} \le \exp(\delta^{-6})$.

The *small item association* of $d \in \{1, 2\}$ and the d-type $\bar{t} \in \mathcal{T}_d$ is the vector $\bar{a}^{d,\bar{t}} \in [0,1]^I$ defined by

$$\bar{a}^{d,\bar{t}}_i = \sum_{C \in \mathcal{C}^* \text{ s.t. } \mathsf{T}^d(C) = \bar{t}} \bar{\lambda}^d_C \cdot C(i), \tag{70}$$

for $i \in I \setminus L$ and $\bar{a}^{d,\bar{t}}_i = 0$ for $i \in L$. Intuitively, $\bar{a}^{d,\bar{t}}_i$ is the fraction of $i \in I \setminus L$ selected by configurations of type $\bar{t}$ in $\bar{\lambda}^d$.

For $d \in \{1,2\}$ define $\bar{v}^d \in [0,1]^I$ by $\bar{v}^d_i = v_d(i)$ for all $i \in I$. Also, we use $\bullet$ to denote element-wise multiplication of two vectors. That is, for $\bar{a}, \bar{b} \in \mathbb{R}^I$ let $\bar{a} \bullet \bar{b} = \bar{c}$, where $\bar{c}_i = \bar{a}_i \cdot \bar{b}_i$ for every $i \in I$. The next lemma will be useful towards adding more vectors to the linear structure.

Lemma 4.23 (Small Items Refinement). *Let $\bar{a} \in [0,1]^I$ be such that* $\mathrm{supp}(\bar{a}) \subseteq I \setminus L$*, let $d \in \{1,2\}$, and let $q \in \mathbb{N}_{\geq 4}$. Then there are subsets $H_1, \ldots, H_q \subseteq I \setminus L$ such that for any $Q \subseteq I \setminus L$ and $\beta \in \left[\frac{1}{q}, 1\right]$ which satisfy*

$$\forall j = 1, \ldots, q: \quad \left\| \mathbb{1}_{Q \cap H_j} \bullet \bar{a} \bullet \bar{v}^d \right\| \leq \beta \left\| \mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d \right\| + \frac{\mathrm{OPT}}{q^5} \max \{ v_d(C \cap H_j) \mid C \in \mathcal{C} \} \tag{71}$$

there is a set $X \subseteq Q$ which admits the following properties:

1. $\left\| \mathbb{1}_X \bullet \bar{a} \bullet (\bar{v}^1 + \bar{v}^2) \right\| \leq \frac{16}{q} \cdot \mathrm{OPT} + 2q\delta$.
2. $\left\| \mathbb{1}_{Q \setminus X} \bullet \bar{a} \bullet \bar{v}^d \right\| \leq \beta \cdot \bar{a} \cdot \bar{v}^d$.

We refer to $H_1, \ldots, H_q$ as the *refinement* of $\bar{a}$ and q in dimension d. Indeed, the condition in (71) is essentially a variant of (45). Lemma 4.23 plays a central role in showing the correctness of the structure $\mathcal{S}$ (see the proof of Lemma 4.29). We defer the proof of Lemma 4.23 to Section 4.2.3.

We select $q = \left\lceil \exp\left(\delta^{-10}\right) \right\rceil$. For any $d, d' \in \{1,2\}$ and $\bar{t} \in \mathcal{T}_d$ let $H^{d,\bar{t},d'}_1, \ldots, H^{d,\bar{t},d'}_q$ be the refinement of $\bar{a}^{d,\bar{t}}$ and q in dimension d'. We use the small items association and its refinement to define additional vectors as follows:

$$\mathcal{S}_{\mathrm{small}} = \left\{ \mathbb{1}_{H^{d,\bar{t},d'}_j} \bullet \bar{a}^{d,\bar{t}} \bullet \bar{v}^{d'} \mid d, d' \in \{1,2\},\ \bar{t} \in \mathcal{T}_d,\ j = 1, \ldots, q \right\} .$$

Finally, the structure is $\mathcal{S} = \mathcal{S}_{\mathrm{large}} \cup \mathcal{S}_{\mathrm{small}}$.

4.2.2 Correctness

We first observe that

$$|\mathcal{S}| = |\mathcal{S}_{\mathrm{large}}| + |\mathcal{S}_{\mathrm{small}}| \leq |\mathcal{G}_1| \cdot |\mathcal{G}_2| + 2 \cdot q \cdot (|\mathcal{T}_1| + |\mathcal{T}_2|) \leq \exp(\delta^{-20}) = \varphi(\delta) .$$

Let $\bar{u} \in \mathcal{S}$ such that $\mathrm{supp}(\bar{u}) \cap L \neq \emptyset$, then $\bar{u} \in \mathcal{S}_{\mathrm{large}}$. Therefore, by (69) there is $(j_1, \ell_1) \in \mathcal{G}_1$ and $(j_2, \ell_2) \in \mathcal{G}_2$ such that $\bar{u} = \mathbb{1}_{G^{1,j_1}_{\ell_1}} \wedge \mathbb{1}_{G^{2,j_2}_{\ell_2}}$. By Definition 4.21, for $d \in \{1,2\}$ there are $q^d_1, q^d_2 \in I_{d,j_d}$ such that $G^{d,j_d}_{\ell_d} = \{ i \in I_{d,j_d} \mid q^d_1 \preceq_{\hat{d}} i \preceq_{\hat{d}} q^d_2 \}$; thus, $\mathbb{1}_{G^{d,j_d}_{\ell_d}} \in \mathcal{S}^*_d$. It follows that $\bar{u} = \mathbb{1}_{G^{1,j_1}_{\ell_1}} \wedge \mathbb{1}_{G^{2,j_2}_{\ell_2}} \in \mathcal{S}^*$.

Let $\beta \in [\delta^5, 1]$ and $\bar{z} \in [0,1]^I$ such that $\bar{z}$ is small-items integral, $\mathrm{supp}(\bar{z}) \subseteq \mathrm{supp}(\bar{w})$, and

$$\bar{z} \cdot \bar{u} \leq \beta \cdot \bar{w} \cdot \bar{u} + \frac{1}{\varphi^{10}(\delta)} \cdot \mathrm{OPT} \cdot \mathsf{tol}(\bar{u}) \tag{72}$$

for all $\bar{u} \in \mathcal{S}$. To verify that $\mathcal{S}$ is a $(\delta, \varphi(\delta))$ linear structure, it remains to show that $\mathrm{OPT}_f(\bar{z}) \leq \beta(1 + 10\delta) \cdot \|\bar{\lambda}\| + \varphi(\delta) + \delta^{10} \cdot \mathrm{OPT}(I, v)$.

We first generate two vectors $\bar{z}^1$ and $\bar{z}^2$ such that $\bar{z} \wedge \mathbb{1}_L$ and $\bar{z}^d \cdot \mathbb{1}_{G^{d,j}_\ell} \lesssim \beta \bar{w}^d \cdot \mathbb{1}_{G^{d,j}_\ell}$ for every $d \in \{1,2\}$ and $(j, \ell) \in \mathcal{G}_d$. Each item $i \in L$ belongs to groups $G^{1,j_1}_{\ell_1}$ and $G^{2,j_2}_{\ell_2}$. The demand $\bar{z}_i$ of i

is partitioned between $\bar{z}^1$ and $\bar{z}^2$ with the same proportion that $\bar{w}^1$ and $\bar{w}^2$ contributed to the total demand of items in $G^{1,j_1}_{\ell_1} \cap G^{2,j_2}_{\ell_2}$. Specifically, for $d \in \{1,2\}$, define $\bar{z}^d \in [0,1]^I$ by

$$\forall (j_1,\ell_1) \in \mathcal{G}_1,\ (j_2,\ell_2) \in \mathcal{G}_2,\ i \in G^{1,j_1}_{\ell_1} \cap G^{2,j_2}_{\ell_2} \cap \mathrm{supp}(\bar{z}): \quad \bar{z}^d_i = \bar{z}_i \cdot \frac{\left(\mathbb{1}_{G^{1,j_1}_{\ell_1}} \wedge \mathbb{1}_{G^{2,j_2}_{\ell_2}}\right) \cdot \bar{w}^d}{\left(\mathbb{1}_{G^{1,j_1}_{\ell_1}} \wedge \mathbb{1}_{G^{2,j_2}_{\ell_2}}\right) \cdot \bar{w}}, \tag{73}$$

and $\bar{z}^d_i = 0$ for any other $i \in I$. Observe that since $\mathrm{supp}(\bar{z}) \subseteq \mathrm{supp}(\bar{w})$ we never get in (73) a division by zero. Since for every $i \in L$ there is a unique $(j_1,\ell_1) \in \mathcal{G}_1$ and a unique $(j_2,\ell_2) \in \mathcal{G}_2$ such that $i \in G^{1,j_1}_{\ell_1} \cap G^{2,j_2}_{\ell_2}$, it follows that $\bar{z} \wedge \mathbb{1}_L = \bar{z}^1 + \bar{z}^2$. For every $d \in \{1,2\}$ and $(j,\ell) \in \mathcal{G}_d$ it holds that

$$\begin{aligned}
\bar{z}^d \cdot \mathbb{1}_{G^{d,j}_\ell} &= \sum_{(j',\ell') \in \mathcal{G}_{\hat{d}}} \ \sum_{i \in G^{d,j}_\ell \cap G^{\hat{d},j'}_{\ell'} \cap \mathrm{supp}(\bar{z})} \bar{z}_i \cdot \frac{\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}^d}{\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}} \\
&= \sum_{(j',\ell') \in \mathcal{G}_{\hat{d}} \text{ s.t. } \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w} \neq 0} \left(\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{z}\right) \cdot \frac{\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}^d}{\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}}.
\end{aligned} \tag{74}$$

Since $\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}} \in \mathcal{S}$, by (72) it holds

$$\begin{aligned}
\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{z} &\leq \beta \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w} + \frac{\mathrm{OPT}}{\varphi^{10}(\delta)} \cdot \mathsf{tol}\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \\
&\leq \beta \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w} + \frac{2 \cdot \delta^{-1}}{\varphi^{10}(\delta)} \cdot \mathrm{OPT}.
\end{aligned} \tag{75}$$

The second inequality holds since there are at most $2\delta^{-1}$ large items in a configuration. Plugging (75) into (74), we have

$$\begin{aligned}
&\bar{z}^d \cdot \mathbb{1}_{G^{d,j}_\ell} \\
\leq &\sum_{(j',\ell') \in \mathcal{G}_{\hat{d}} \text{ s.t. } \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w} \neq 0} \left(\beta \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w} + \frac{2 \cdot \delta^{-1}\mathrm{OPT}}{\varphi^{10}(\delta)}\right) \cdot \frac{\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}^d}{\left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}} \\
\leq &\sum_{(j',\ell') \in \mathcal{G}_{\hat{d}} \text{ s.t. } \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w} \neq 0} \beta \left(\mathbb{1}_{G^{d,j}_\ell} \wedge \mathbb{1}_{G^{\hat{d},j'}_{\ell'}}\right) \cdot \bar{w}^d \ + \ \frac{\delta^{-6}}{\varphi^{10}(\delta)} \cdot \mathrm{OPT} \\
\leq &\beta \cdot \mathbb{1}_{G^{d,j}_\ell} \cdot \bar{w}^d + \frac{\delta^{-6}}{\varphi^{10}(\delta)} \cdot \mathrm{OPT},
\end{aligned} \tag{76}$$

where the second inequality holds since $|\mathcal{G}_{\hat{d}}| \leq 2 \cdot \delta^{-4}$.

Therefore, for every $d \in \{1,2\}$ there is a vector $\bar{r}^d \in [0,1]^I$ such that, for any $(j,\ell) \in \mathcal{G}_d$,

$$\left(\bar{z}^d - \bar{r}^d\right) \cdot \mathbb{1}_{G^{d,j}_\ell} \leq \max\left\{\beta \cdot \mathbb{1}_{G^{d,j}_\ell} \cdot \bar{w}^d - 2, 0\right\}, \tag{77}$$

for every $i \in I$ it holds that $r_i^d \leq z_i^d$, and $\|\bar{r}^d\| \leq \left(2 + \frac{\delta^{-6}}{\varphi^{10}(\delta)} \cdot \text{OPT}\right) \cdot |\mathcal{G}_d| \leq \delta^{-5} + \frac{\delta^{-11}}{\varphi^{10}(\delta)}\text{OPT}$. Hence, $\text{OPT}_f(\bar{r}^d) \leq \|\bar{r}^d\| \leq \delta^{-5} + \frac{\delta^{-11}}{\varphi^{10}(\delta)}\text{OPT}$, as $\sum_{i \in I} \bar{r}_i^d \cdot \mathbb{1}_{\{i\}}$ is a solution for $\text{LP}(\bar{r}^d)$.

For any $d \in \{1,2\}$, let $F_d = \bigcup_{j \in [2h] \text{ s.t. } (j,1) \in \mathcal{G}_d} G_1^{d,j}$ be the set of all items that belong to a first group in one of the fractional groupings $G_1^{d,j}, \ldots, G_{\tau_{d,j}}^{d,j}$. By (77),

$$\begin{aligned}(\bar{z}^d - \bar{r}^d) \cdot \mathbb{1}_{F_d} &\leq \sum_{j \in \{1,\ldots,2h\} \text{ s.t. } (j,1) \in \mathcal{G}_d} (\bar{z}^d - \bar{r}^d) \cdot \mathbb{1}_{G_1^{d,j}} \leq \sum_{j \in \{1,\ldots,2h\} \text{ s.t. } (j,1) \in \mathcal{G}_d} \max\left\{\beta \cdot \mathbb{1}_{G_1^{d,j}} \cdot \bar{w}^d - 2, 0\right\} \\ &\leq \beta \sum_{j \in \{1,\ldots,2h\} \text{ s.t. } (j,1) \in \mathcal{G}_d} \frac{\mathbb{1}_{I_{d,j}} \cdot \bar{w}^d}{h} = \beta \frac{\bar{w}^d \cdot \mathbb{1}_L}{h} \leq 2 \cdot \beta \cdot \delta \cdot \|\bar{\lambda}^d\|\end{aligned}$$

where the third inequality is by Definition 4.21, and the last inequality follows from $h = \delta^{-2}$ and

$$\sum_{i \in L} \bar{w}_i^d = \sum_{C \in \mathcal{C}^*} \bar{\lambda}_C^d \cdot \sum_{i \in L} C(i) \leq \sum_{C \in \mathcal{C}^*} \bar{\lambda}_C^d \cdot 2\delta^{-1} = 2 \cdot \delta^{-1} \|\bar{\lambda}^d\| \ .$$

Define $Q = \text{supp}(\bar{z}) \setminus L = \{i \in I \setminus L \mid \bar{z}_i = 1\}$ and

$$\bar{y} = \sum_{d \in \{1,2\}} \left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right) + \mathbb{1}_Q \ . \tag{78}$$

Then,

$$\begin{aligned}\text{OPT}_f(\bar{z}) &\leq \sum_{d \in \{1,2\}} \left(\text{OPT}_f(\bar{r}^d) + \text{OPT}_f((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{F_d})\right) + \text{OPT}_f(\bar{y}) \\ &\leq \text{OPT}_f(\bar{y}) \ + 2\beta\delta\|\bar{\lambda}\| + 2\delta^{-5} + \frac{2 \cdot \delta^{-11}}{\varphi^{10}(\delta)}\text{OPT} \ .\end{aligned} \tag{79}$$

We proceed to derive an upper bound on $\text{OPT}_f(\bar{y})$, which in turn implies an upper bound on $\text{OPT}_f(\bar{z})$.

Given $d \in \{1,2\}$ we define the *d-size* of $(j,\ell) \in \mathcal{G}_d$, denoted $s^d(j,\ell) \in [0,1]^2$, by $s_d^d(j,\ell) = \frac{\delta^2}{2} j$ and $s_{\hat{d}}^d = \min\{v_{\hat{d}}(i) \mid i \in G_\ell^{d,j}\}$. The value $s^d(j,\ell)$ can be viewed a the rounded volume of items in $G_\ell^{d,j}$.

The next lemma gives the basis for our *shifting* argument.

Lemma 4.24. *Let $d \in \{1,2\}$, $(j,\ell) \in \mathcal{G}_d$ and $i \in G_\ell^{d,j}$. If $\ell \neq 1$ then $v(i) \leq s^d(j,\ell-1)$.*

Proof. As $i \in G_\ell^{d,j} \subseteq I_{d,j}$, it follows that $v_d(i) \leq \frac{\delta^2}{2} \cdot j = s_d^d(j,\ell-1)$. Furthermore, $v_{\hat{d}}(i') \geq v_{\hat{d}}(i)$ for every $i' \in G_{\ell-1}^{d,j}$ as $(G_{\ell'}^{d,j})_{\ell'=1}^{\tau_{d,j}}$ is an h-fractional grouping with respect to the relation $\succeq_{\hat{d}}$. Hence,

$$v_{\hat{d}}(i) \leq \min\left\{v_{\hat{d}}(i') \mid i' \in G_{\ell-1}^{d,j}\right\} = s_{\hat{d}}^d(j,\ell-1) \ .$$

□

We extend the definition of size to d-types by $s^d(\bar{t}) = \sum_{(j,\ell) \in \mathcal{G}_d} \bar{t}_{(j,\ell)} \cdot s^d(j,\ell)$ for any $d \in \{1,2\}$ and $\bar{t} \in \mathcal{T}_d$.

Lemma 4.25. *Let $d \in \{1,2\}$ and $C \in \mathcal{C}^*$ with $\bar{\lambda}_C^d > 0$. Then $\sum_{i \in I \setminus L} v(i) \cdot C(i) \leq \mathbf{1} - s^d\left(\mathsf{T}^d(C)\right)$.*

Proof. For any $i \in L$ such that $C(i) > 0$ there is a unique $(j,\ell) \in \mathcal{G}_d$ for which $i \in G_\ell^{d,j}$. Thus,

$$\sum_{i \in I \setminus L} v(i) \cdot C(i) = \sum_{i \in I} v(i) \cdot C(i) - \sum_{i \in L} v(i) \cdot C(i) = v(C) - \sum_{(j,\ell) \in \mathcal{G}_d} \sum_{i \in G_\ell^{d,j}} v(i) \cdot C(i) \ . \tag{80}$$

Therefore, we have

$$\begin{aligned}\sum_{i\in I\setminus L} v_d(i)\cdot C(i) &= v_d(C) - \sum_{(j,\ell)\in\mathcal{G}_d}\sum_{i\in G_\ell^{d,j}} v_d(i)\cdot C(i)\\ &\leq 1-\delta-\sum_{(j,\ell)\in\mathcal{G}_d}\sum_{i\in G_\ell^{d,j}}\left(s_d^d(j,\ell)-\frac{\delta^2}{2}\right)\cdot C(i)\\ &= 1-\delta-\sum_{(j,\ell)\in\mathcal{G}_d}\sum_{i\in G_\ell^{d,j}} s_d^d(j,\ell)\cdot C(i)+\frac{\delta^2}{2}\sum_{(j,\ell)\in\mathcal{G}_d}\sum_{i\in G_\ell^{d,j}}\cdot C(i)\\ &= 1-\delta-\sum_{(j,\ell)\in\mathcal{G}_d}\mathsf{T}^d_{(j,\ell)}(C)\cdot s_d^d(j,\ell)+\frac{\delta^2}{2}\sum_{i\in L} C(i)\\ &\leq 1-s_d^d(\mathsf{T}^d(C))\ .\end{aligned}\tag{81}$$

The first equality is by (80). The first inequality holds, as C has δ-slack in dimension d since $\bar{\lambda}_C^d>0$, and since $v_d(i)>\frac{\delta^2}{2}(j-1)$ for any $i\in G_\ell^{d,j}\subseteq I_{d,j}$. The last inequality holds as there are at most $2\delta^{-1}$ large items in a multi-configuration. Similarly,

$$\begin{aligned}\sum_{i\in I\setminus L} v_{\hat{d}}(i)\cdot C(i) &= v_{\hat{d}}(C)-\sum_{(j,\ell)\in\mathcal{G}_d}\sum_{i\in G_\ell^{d,j}} v_{\hat{d}}(i)\cdot C(i)\\ &\leq 1-\sum_{(j,\ell)\in\mathcal{G}_d}\sum_{i\in G_\ell^{d,j}} s_{\hat{d}}^d(j,\ell)\cdot C(i)\\ &= 1-\sum_{(j,\ell)\in\mathcal{G}_d}\mathsf{T}^d_{(j,\ell)}(C)\cdot s_{\hat{d}}^d(j,\ell)\\ &\leq 1-s_{\hat{d}}^d(\mathsf{T}^d(C))\ .\end{aligned}\tag{82}$$

The first equality follows from (80) and the first inequality is by the definition of $s_{\hat{d}}^d(j,\ell)$. The statement of the lemma follows from (81) and (82). $\square$

For any $d\in\{1,2\}$ and $\bar{t}\in\mathcal{T}_d$, the *prevalence* of type $\bar{t}$ is $\eta_d(\bar{t})=\sum_{C\in\mathcal{C}^* \text{ s.t. } \mathsf{T}^d(C)=\bar{t}}\bar{\lambda}_C^d$. Informally, $\eta_d(\bar{t})$ is the number of configurations of type $\bar{t}$ selected by $\bar{\lambda}^d$. Also, define $\kappa_d(\bar{t})=\lceil\beta\cdot\eta_d(\bar{t})\rceil+2\cdot\delta^{-1}$ for any $d\in\{1,2\}$ and $\bar{t}\in\mathcal{T}_d$. We construct a solution of LP$(\bar{y})$ in which there are $\kappa_d(\bar{t})$ configurations with large items of total size at most $s^d(\bar{t})$. For the assignment of large items we use the next lemma.

Lemma 4.26. *There are vectors $\bar{x}^{d,\bar{t}}\in[0,1]^{\mathcal{C}}$ for $d\in\{1,2\}$ and $\bar{t}\in\mathcal{T}_d$ such that*

1. *for any $d\in\{1,2\}$ the coverage of $\sum_{\bar{t}\in\mathcal{T}_d}\kappa_d(\bar{t})\cdot\bar{x}^{d,\bar{t}}$ is $\left(\bar{z}^d-\bar{r}^d\right)\wedge\mathbb{1}_{L\setminus F_d}$,*
2. *for any $d\in\{1,2\}$ and $\bar{t}\in\mathcal{T}_d$ it holds that $\|\bar{x}^{d,t}\|=1$,*
3. *and for any $d\in\{1,2\}$, $\bar{t}\in\mathcal{T}_d$ and $C\in\operatorname{supp}(\bar{x}^{d,\bar{t}})$, it holds that $v(C)\leq s^d(\bar{t})$.*

The proof of Lemma 4.26 relies on the following combinatorial claim (we omit the proof).

Claim 4.27. *Let E be an arbitrary finite set, $\xi\in\mathbb{N}_+$ and $\bar{\gamma}\in\left[0,\frac{1}{\xi}\right]^E$ such that $\|\bar{\gamma}\|\leq 1$. Then there exists a random set $K\subseteq E$ such that $|K|\leq\xi$ and $\Pr(e\in K)=\xi\cdot\bar{\gamma}_e$ for every $e\in E$.*

Proof of Lemma 4.26. Let $d \in \{1,2\}$ and for any $(j,\ell) \in \mathcal{G}_d$, define $\rho_{(j,\ell)} = \sum_{\bar{t} \in \mathcal{T}_d} \bar{t}_{(j,\ell)} \cdot \kappa_d(\bar{t})$. Then $\rho_{(j,\ell)} \geq 2 \cdot \delta^{-1}$. For any $(j,\ell) \in \mathcal{G}_d$ and $i \in G^{d,j}_{\ell}$ such that $\ell \neq 1$, define $p_i = \frac{\bar{z}^d_i - \bar{r}^d_i}{\rho_{(j,\ell-1)}} \leq \frac{1}{2 \cdot \delta^{-1}}$.

For every $(j,\ell) \in \mathcal{G}_d$ with $\ell \neq 1$ it holds that

$$
\begin{aligned}
\rho_{(j,\ell-1)} &= \sum_{\bar{t} \in \mathcal{T}_d} \bar{t}_{(j,\ell-1)} \cdot \kappa_d(\bar{t}) \\
&\geq \beta \sum_{\bar{t} \in \mathcal{T}_d} \bar{t}_{(j,\ell-1)} \cdot \eta_d(\bar{t}) = \beta \mathbb{1}_{G^{d,j}_{\ell-1}} \cdot \bar{w}^d \\
&\geq \beta \frac{\bar{w}^d \cdot \mathbb{1}_{I_{d,j}}}{h} \\
&\geq \max\left\{\beta \cdot \bar{w}^d \cdot \mathbb{1}_{G^{d,j}_{\ell}} - 1,\ 0\right\} \geq \left(\bar{z}^d - \bar{r}^d\right) \cdot \mathbb{1}_{G^{d,j}_{\ell}} \ .
\end{aligned}
$$

The second and third inequalities hold since $G^{d,j}_1, \ldots, G^{d,j}_{\tau_{d,j}}$ is an h-fractional grouping of $I_{d,j}$. The last inequality is by (77). Therefore, $\sum_{i \in G^{d,j}_{\ell}} p_i \leq 1$.

Fix $\bar{t} \in \mathcal{T}_d$, and for any $(j,\ell) \in \mathcal{G}_d$ with $\ell \neq 1$ let $K_{(j,\ell)} \subseteq G^{d,j}_{\ell}$ be a random set such that $|K_{(j,\ell)}| \leq \bar{t}_{(j,\ell-1)}$ and $\Pr(i \in K_{(j,\ell)}) = \bar{t}_{(j,\ell-1)} \cdot p_i$ for every $i \in G^{d,j}_{\ell}$. The random sets $K_{(j,\ell)}$ exist by Claim 4.27. Furthermore, we may assume the random sets $\left(K_{(j,\ell)}\right)_{(j,\ell) \in \mathcal{G}_d,\ \ell \neq 1}$ are independent. Define $R = \bigcup_{(j,\ell) \in \mathcal{G}_d \text{ s.t. } \ell \neq 1} K_{(j,\ell)}$ and $\bar{x}^{d,\bar{t}}_C = \Pr(R = C)$ for all $C \in \mathcal{C}$. It follows that $\|\bar{x}^{d,\bar{t}}\| = \sum_{C \in \mathcal{C}^*} \Pr(R = C) = 1$. Observe that

$$
v(R) \leq \sum_{(j,\ell) \in \mathcal{G}_d \text{ s.t. } \ell \neq 1} v(K_{(j,\ell)}) \leq \sum_{(j,\ell) \in \mathcal{G}_d \text{ s.t. } \ell \neq 1} \bar{t}_{(j,\ell-1)} \cdot s^d(j,\ell-1) \leq s^d(\bar{t}) \ .
$$

The second inequality holds since $|K_{(j,\ell)}| \leq \bar{t}_{(j,\ell-1)}$ and for every $i \in K_{(j,\ell)}$ it holds that $v(i) \leq s^d(j,\ell-1)$ by Lemma 4.25. Thus, for every $C \in \text{supp}(\bar{x}^{d,\bar{t}})$ we have that $v(C) \leq s^d(\bar{t})$. Finally, for every $i \in \text{supp}\left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right)$, there is $(j,\ell) \in \mathcal{G}_d$ with $\ell \neq 1$ such that $i \in G^{d,j}_{\ell}$. Hence,

$$
\sum_{C \in \mathcal{C}} \bar{x}^{d,\bar{t}}_C \cdot C(i) = \Pr(i \in R) = \bar{t}_{(j,\ell-1)} \cdot \frac{\bar{z}^d_i - \bar{r}^d_i}{\rho_{(j,\ell-1)}} \ . \tag{83}
$$

Let $\bar{w}'$ be the coverage of $\sum_{\bar{t} \in \mathcal{T}_d} \kappa_d(\bar{t}) \cdot \bar{x}^{d,\bar{t}}$. By construction, we have $\bar{w}'_i = 0$ for any $i \in I$ such that $i \notin \text{supp}\left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right)$. For any $i \in \text{supp}\left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right)$, it holds that

$$
\bar{w}'_i = \sum_{C \in \mathcal{C}} \sum_{\bar{t} \in \mathcal{T}_d} \kappa_d(\bar{t}) \cdot \bar{x}^{d,\bar{t}}_C \cdot C(i) = \sum_{\bar{t} \in \mathcal{T}_d} \kappa_d(\bar{t}) \cdot \bar{t}_{(j,\ell-1)} \cdot \frac{\bar{z}^d_i - \bar{r}^d_i}{\rho_{(j,\ell-1)}} = \bar{z}^d_i - \bar{r}^d_i,
$$

where the second equality is by (83), and the last equality is by the definition of $\rho_{(j,\ell)}$. $\square$

Recall that $Q = \text{supp}(\bar{z}) \setminus L$. The assignment of items in Q relies on integrality properties of polytopes. Define $M = \exp(-\delta^{-9}) \cdot \text{OPT} + \exp(\delta^{-11})$ and

$$
B = \{(d, \bar{t}, m) \mid d \in \{1,2\},\ \bar{t} \in \mathcal{T}_d,\ m \in [\kappa_d(\bar{t})]\} \cup \{1, \ldots, M\} \ .
$$

We consider B as a set of bins, and define a polytope

$$P = \left\{ \bar{\mu} \in [0,1]^{Q\times B} \;\middle|\; \begin{array}{ll} \sum_{b\in B} \bar{\mu}_{i,b} = 1 & \forall i \in Q \\ \sum_{i\in Q} \bar{\mu}_{i,(d,\bar{t},m)} \cdot v(i) \leq \mathbf{1} - s^d(\bar{t}) & \forall d \in \{1,2\},\ \bar{t}\in\mathcal{T}_d,\ m\in\{1,\ldots,\kappa_d(\bar{t})\} \\ \sum_{i\in Q} \bar{\mu}_{i,m}\cdot v(i) \leq \mathbf{1} & \forall m \in \{1,\ldots,M\} \end{array} \right\} \tag{84}$$

The entry $\bar{\mu}_{i,b}$ in P represents a fractional assignment of an item $i \in Q$ to bin b. The first constraint in (84) represents the requirement that each item is fully assigned, and the remaining constraints represent a volume limit for each bin.

The following is a well known integrality property of P (see, e.g., Bansal et al. [BEK16]).

Lemma 4.28. *Let $\bar{\mu}$ be a vertex of P. Then $|\{i \in Q \mid \exists b \in B:\ 0 < \bar{\mu}_{i,b} < 1\}| \leq 2\cdot|B|$.*

Before we use Lemma 4.28, we need to show that P has a vertex.

Lemma 4.29. *It holds $P \neq \emptyset$.*

Proof. Ideally, we would like to define $\bar{\mu}_{i,(d,\bar{t},m)} = \frac{a_i^{d,\bar{t}}}{\kappa_d(\bar{t})}$ for any $i \in Q$, $d \in \{1,2\}$, $\bar{t}\in\mathcal{T}_d$ and $m \in \{1,\ldots,\kappa_d(\bar{t})\}$. Using (72) we can show that $\sum_{i\in Q}\bar{\mu}_{i,(d,\bar{t},m)}\cdot v_{d'}(i)$ is not significantly larger than $\mathbf{1}-s^d(\bar{t})$; however, we cannot show it is smaller (or equal) to $\mathbf{1}-s^d(\bar{t})$. Thus, the suggested vector $\bar{\mu}$ may not satisfy the properties in (84). We use Lemma 4.23 to overcome this difficulty. Specifically, we define $\bar{\mu}_{i,(d,\bar{t},m)} = \frac{a_i^{d,\bar{t}}}{\kappa_d(\bar{t})}$ for items $i \in Q\setminus X_1\setminus X_2$, where the sets X_1 and X_2 are obtained via Lemma 4.23. The value of $\bar{\mu}_{i,m}$ is subsequently increased for $i \in X_1\cup X_2$ to ensure the first constraint in (84) holds. Property 1 of Lemma 4.23 is used to show that $\sum_{i\in Q}\bar{\mu}_{i,m}\cdot v(i)\leq \mathbf{1}$, and property 2 of the lemma is used to show that $\sum_{i\in Q}\bar{\mu}_{i,(d,\bar{t},m)}\cdot v(i)\leq \mathbf{1}-s^d(\bar{t})$. We now proceed to the formal proof.

Recall that $H_1^{d,\bar{t},d'},\ldots,H_q^{d,\bar{t},d'}$ is the refinement of $\bar{a}^{d,\bar{t}}$ and $q = \lceil \exp(\delta^{-10})\rceil$ in dimension d'. For every $d,d'\in\{1,2\}$, $\bar{t}\in\mathcal{T}_d$ and $j = 1,\ldots,q$ it holds that

$$\begin{aligned}\sum_{i\in H_j^{d,\bar{t},d'}\cap Q} \bar{a}_i^{d,\bar{t}}\cdot v_{d'}(i) &= \bar{z}\cdot\left(\mathbb{1}_{H_j^{d,\bar{t},d'}}\bullet \bar{a}^{d,\bar{t}}\bullet \bar{v}^{d'}\right)\\ &\leq \beta\cdot\bar{w}\cdot\left(\mathbb{1}_{H_j^{d,\bar{t},d'}}\bullet \bar{a}^{d,\bar{t}}\bullet \bar{v}^{d'}\right) + \frac{1}{\varphi^{10}(\delta)}\cdot \mathrm{OPT}\cdot\max\left\{\sum_{i\in C}\mathbb{1}_{i\in H_j^{d,\bar{t},d'}}\cdot \bar{a}_i^{d,\bar{t}}\cdot \bar{v}_{d'}(i)\;\middle|\; C\in\mathcal{C}\right\}\\ &\leq \beta\cdot\|\mathbb{1}_{H_j^{d,\bar{t},d'}}\bullet \bar{a}^{d,\bar{t}}\bullet \bar{v}^{d'}\| + \frac{1}{\varphi^{10}(\delta)}\cdot\mathrm{OPT}\cdot\max\left\{v_{d'}(H_j^{d,\bar{t},d'}\cap C)\;\middle|\; C\in\mathcal{C}\right\}\ .\end{aligned}$$

The equality follows from the definition of Q. The first inequality follows from (72) and the fact that $\mathbb{1}_{H_j^{d,\bar{t},d'}}\bullet\bar{a}^{d,\bar{t}}\bullet\bar{v}^{d'}\in\mathcal{S}_{\text{small}}\subseteq\mathcal{S}$. The second inequality holds, as $\bar{w}$ is small-items integral and $\mathrm{supp}(\bar{a}^{d,\bar{t}})\subseteq\mathrm{supp}(\bar{w})\setminus L$. Thus, by Lemma 4.23, for every $d,d'\in\{1,2\}$, $\bar{t}\in\mathcal{T}_d$ and $j=1,\ldots,q$ there is a set $X^{d,\bar{t},d'}\subseteq Q$ such that

$$\left\|\mathbb{1}_{X^{d,\bar{t},d'}}\bullet\bar{a}^{d,\bar{t}}\bullet(\bar{v}^1+\bar{v}^2)\right\|\leq\frac{16}{q}\cdot\mathrm{OPT}+2q\delta \quad\text{and}\quad \left\|\mathbb{1}_{Q\setminus X^{d,\bar{t},d'}}\bullet\bar{a}^{d,\bar{t}}\bullet\bar{v}^{d'}\right\|\leq\beta\cdot\bar{a}^{d,\bar{t}}\cdot\bar{v}^{d'}\ . \tag{85}$$

Define $\bar{\mu} \in [0,1]^{Q\times B}$ by

$$\bar{\mu}_{i,(d,\bar{t},m)} = \begin{cases} \frac{\bar{a}_i^{d,\bar{t}}}{\kappa_d(\bar{t})}, & i \in Q \setminus X^{d,\bar{t},1} \setminus X^{d,\bar{t},2}, \\ 0, & \text{otherwise} \ . \end{cases}$$

for every $i \in Q$, $d \in \{1,2\}$, $\bar{t} \in \mathcal{T}_d$ and $m = 1,\ldots,\kappa_d(\bar{t})$. Also, for every $i \in Q$ and $m = 1,\ldots,M$ define

$$\bar{\mu}_{i,m} = \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \frac{\bar{a}_i^{d,\bar{t}} \cdot \mathbb{1}_{i\in X^{d,\bar{t},1}\cup X^{d,\bar{t},2}}}{M} \ .$$

Next, we show that $\bar{\mu} \in P$. For every $i \in Q$ it holds that

$$\begin{aligned}
\sum_{b\in B} \bar{\mu}_{i,b} &= \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \sum_{m\in[\kappa_d(\bar{t})]} \bar{\mu}_{i,(d,\bar{t},m)} + \sum_{m\in[M]} \bar{\mu}_{i,m} \\
&= \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \kappa_d(\bar{t}) \cdot \frac{\bar{a}_i^{d,\bar{t}}}{\kappa_d(\bar{t})} \cdot \mathbb{1}_{i\in Q\setminus X^{d,\bar{t},1}\setminus X^{d,\bar{t},2}} + \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} M \cdot \frac{a_i^{d,\bar{t}}}{M} \cdot \mathbb{1}_{i\in X^{d,\bar{t},1}\cup X^{d,\bar{t},2}} \\
&= \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \bar{a}^{d,\bar{t}} \cdot \mathbb{1}_{i\in Q\setminus X^{d,\bar{t},1}\setminus X^{d,\bar{t},2}} + \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} a^{d,\bar{t}} \cdot \mathbb{1}_{i\in X^{d,\bar{t},1}\cup X^{d,\bar{t},2}} \\
&= \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \bar{a}^{d,\bar{t}} \\
&= \bar{w}_i^1 + \bar{w}_i^2 = 1,
\end{aligned}$$

where the fifth equality follows from (70).

For every $d, d' \in \{1,2\}$, $\bar{t} \in \mathcal{T}_d$ we have

$$\begin{aligned}
\bar{a}^{d,\bar{t}} \cdot \bar{v}^{d'} &= \sum_{i\in I\setminus L} v_{d'}(i) \sum_{C\in\mathcal{C}^* \text{ s.t } \mathsf{T}^d(C)=\bar{t}} \bar{\lambda}_C^d \cdot C(i) = \sum_{C\in\mathcal{C}^* \text{ s.t } \mathsf{T}^d(C)=\bar{t}} \bar{\lambda}_C^d \cdot \sum_{i\in I\setminus L} v_{d'}(i) \cdot C(i) \\
&\leq \sum_{C\in\mathcal{C}^* \text{ s.t } \mathsf{T}^d(C)=\bar{t}} \bar{\lambda}_C^d \cdot \left(1 - s_{d'}^d(\bar{t})\right) = \left(1 - s_{d'}^d(\bar{t})\right) \cdot \eta_d(\bar{t}),
\end{aligned}$$

where the first equality is by (70) and the inequality is by Lemma 4.25. Thus, for $m = 1,\ldots,\kappa_d(\bar{t})$ we have

$$\sum_{i\in Q} \bar{\mu}_{i,(d,t,m)} \cdot v_{d'}(i) = \sum_{i\in Q\setminus X^{d,\bar{t},1}\setminus X^{d,\bar{t},2}} \frac{\bar{a}_i^{d,\bar{t}} \cdot v_{d'}(i)}{\kappa_d(\bar{t})} \leq \frac{\beta \cdot \bar{a}^{d,\bar{t}} \cdot \bar{v}^{d'}}{\kappa_d(\bar{t})} \leq \frac{\beta \cdot \left(1 - s_{d'}^d(\bar{t})\right) \eta_d(\bar{t})}{\kappa_d(\bar{t})} \leq 1 - s_{d'}^d(\bar{t}),$$

where the first inequality is by (85).

Finally, for every $m = 1,\ldots,M$ and $d' \in \{1,2\}$ we have

$$\begin{aligned}
\sum_{i\in Q} \bar{\mu}_{i,m} \cdot v_{d'}(i) &= \sum_{i\in Q} v_{d'}(i) \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \frac{\bar{a}_i^{d,\bar{t}} \cdot \mathbb{1}_{i\in X^{d,\bar{t},1}\cup X^{d,\bar{t},2}}}{M} \\
&\leq \frac{1}{M} \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \left(\|\mathbb{1}_{X^{d,\bar{t},1}} \bullet \bar{a}^{d,\bar{t}} \bullet \bar{v}^{d'}\| + \|\mathbb{1}_{X^{d,\bar{t},2}} \bullet \bar{a}^{d,\bar{t}} \bullet \bar{v}^{d'}\| \right) \\
&\leq \frac{1}{M} \sum_{d\in\{1,2\}} \sum_{\bar{t}\in\mathcal{T}_d} \left(\frac{32}{q} \cdot \mathrm{OPT} + 4q\delta \right) \leq 1,
\end{aligned}$$

where the second inequality is by (85) and the last inequality holds since $|\mathcal{T}_d| \leq \exp(\delta^{-6})$, $q \geq \exp(\delta^{-10})$ and $M = \exp(-\delta^{-9}) \cdot \mathrm{OPT} + \exp(\delta^{-11})$. Thus, $\bar{\mu} \in P$, i.e., $P \neq \emptyset$. $\square$

We now have the tools to prove the following.

Lemma 4.30. *It holds that* $\mathrm{OPT}_f(\bar{y}) \leq (1+8\delta)|B| + 1$.

Proof. Let $\bar{\mu}^*$ be a vertex of P, and let $Q_I = \{i \in Q \mid \exists b \in B : \ \bar{\mu}^*_{i,b} = 1\}$. By Lemma 4.28 it holds that $|Q \setminus Q_I| \leq 2|B|$. As $Q \subseteq I \setminus L$, it follows that the items of $Q \setminus Q_I$ can be packed into $4\delta|Q \setminus Q_I| + 1 \leq 8\delta|B| + 1$ bins using the First-Fit strategy (Lemma 2.5). Thus, $\mathrm{OPT}_f(\mathbb{1}_{Q \setminus Q_I}) \leq 8\delta|B| + 1$.

For every $b \in B$ define $C_b = \{i \in Q \mid \bar{\mu}^*_i = 1\}$. It follows that $Q_I = \bigcup_{b \in B} C_b$. Recall that $\bar{x}^{d,\bar{t}}$ are the vectors defined in Lemma 4.26. For every $(d, \bar{t}, m) \in B \setminus \{1, \ldots, M\}$ define a vector $\bar{\gamma}^{d,\bar{t},m} \in [0,1]^{\mathcal{C}}$ by $\bar{\gamma}^{d,\bar{t},m}_{C \cup C_{d,\bar{t},m}} = \bar{x}^{d,\bar{t}}_C$ for any $C \in \mathrm{supp}(\bar{x}^{d,\bar{t}})$, and $\bar{\gamma}^{d,\bar{t},m}_{C'} = 0$ for any other configuration $C' \in \mathcal{C}$. By definition of P, it holds that $v(C_{d,\bar{t},m}) \leq \mathbf{1} - s^d(\bar{t})$, and by Lemma 4.26, for every $C \in \mathrm{supp}(\bar{x}^{d,\bar{t}})$ it holds that $v(C) \leq s^d(\bar{t})$; thus, $C \cup C_{d,\bar{t},m} \in \mathcal{C}$, and $\bar{\gamma}^{d,\bar{t},m}$ is well defined. Also, for any $m = 1, \ldots, M$ define $\bar{\gamma}^m \in [0,1]^{\mathcal{C}}$ by $\bar{\gamma}^m_{C_m} = 1$ and $\bar{\gamma}^m_C = 0$ for $C \in \mathcal{C} \setminus \{C_m\}$.

Define $\bar{x} = \sum_{b \in B} \bar{\gamma}^b$. We show that $\bar{x}$ is a solution for LP $\left(\sum_{d \in \{1,2\}} \left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right) + \mathbb{1}_{Q_I}\right)$. For $i \in L$ we have

$$
\begin{aligned}
\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) &= \sum_{C \in \mathcal{C}} \sum_{b \in B} \bar{\gamma}^b_C \cdot C(i) \\
&= \sum_{C \in \mathcal{C}} \sum_{d \in \{1,2\}} \sum_{\bar{t} \in \mathcal{T}_d} \sum_{m \in [\kappa_d(\bar{t})]} \bar{x}^{d,\bar{t}}_C \cdot C(i) \\
&= \sum_{d \in \{1,2\}} \sum_{C \in \mathcal{C}} \sum_{\bar{t} \in \mathcal{T}_d} \kappa_d(\bar{t}) \cdot \bar{x}^{d,\bar{t}}_C \cdot C(i) \\
&= \sum_{d \in \{1,2\}} \left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right) \ .
\end{aligned}
$$

The second equality holds by definition of $\bar{\gamma}^b$, and since the sets C_b do not contain large items. The last equality is by Lemma 4.26. For any $i \in Q_I$ there is a unique $b \in B$ such that $i \in C_b$. Thus, $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = \sum_{C \in \mathcal{C}} \bar{\gamma}^b_C \cdot C(i) = 1$. Therefore, $\bar{x}$ is a solution for the linear program LP $\left(\sum_{d \in \{1,2\}} \left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right) + \mathbb{1}_{Q_I}\right)$. As $\|\bar{\gamma}^b\| = 1$ for every $b \in B$, it follows that $\|\bar{x}\| = B$. Thus,

$$
\mathrm{OPT}_f\left(\sum_{d \in \{1,2\}} \left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right) + \mathbb{1}_{Q_I}\right) \leq \|\bar{x}\| = B,
$$

and by definition of $\bar{y}$ (78), we have

$$
\mathrm{OPT}_f(\bar{y}) = \mathrm{OPT}_f\left(\sum_{d \in \{1,2\}} \left((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}\right) + \mathbb{1}_{Q_I}\right) + \mathrm{OPT}_f(\mathbb{1}_{Q \setminus Q_I}) \leq (1+8\delta)|B| + 1 \ . \quad \square
$$

Observe that

$$
\begin{aligned}
|B| = \sum_{d \in \{1,2\}} \sum_{\bar{t} \in \mathcal{T}_d} \kappa_d(t) + M &= \sum_{d \in \{1,2\}} \sum_{\bar{t} \in \mathcal{T}_d} \left(\lceil \beta \eta_d(t) \rceil + 2\delta^{-1}\right) + \exp(-\delta^{-9}) \cdot \mathrm{OPT} + \exp(\delta^{-11}) \\
&\leq \beta\|\bar{\lambda}\| + (|\mathcal{T}_1| + |\mathcal{T}_2|) \cdot (1 + 2\delta^{-1}) + \exp(-\delta^{-9}) \cdot \mathrm{OPT} + \exp(\delta^{-11}) \\
&\leq \beta\|\bar{\lambda}\| + \exp(-\delta^{-9}) \cdot \mathrm{OPT} + \exp(\delta^{-12}) \ .
\end{aligned} \tag{86}
$$

The first inequality holds since $\sum_{\bar{t}\in\mathcal{T}_d}\eta_d(\bar{t})=\|\bar{\lambda}^d\|$, and the second inequality uses $|\mathcal{T}_d|\leq\exp(\delta^{-6})$. By (79) we have

$$\begin{aligned}
\mathrm{OPT}_f(\bar{z}) &\leq \mathrm{OPT}_f(\bar{y}) \ +2\beta\delta\|\bar{\lambda}\|+2\delta^{-5}+\frac{2\cdot\delta^{-11}}{\varphi^{10}(\delta)}\mathrm{OPT}\\
&\leq (1+8\delta)|B|+1+2\beta\delta\|\bar{\lambda}\|+2\delta^{-5}+\frac{2\delta^{-11}}{\varphi^{10}(\delta)}\mathrm{OPT}\\
&\leq (1+8\delta)\left(\beta\|\bar{\lambda}\|+\exp(-\delta^{-9})\cdot\mathrm{OPT}+\exp(\delta^{-12})\right)+1+2\delta\beta\|\bar{\lambda}\|+2\delta^{-5}+\frac{2\cdot\delta^{-11}}{\varphi^{10}(\delta)}\mathrm{OPT}\\
&\leq \beta(1+10\delta)\|\bar{\lambda}\|+\exp(\delta^{-20})+\delta^{10}\mathrm{OPT},
\end{aligned}$$

where the second inequality is by Lemma 4.30, the third inequality is by (86), and the last inequality uses $\varphi(\delta)=\exp(\delta^{-20})$. Thus, we showed that $\mathcal{S}$ is a linear structure, which completes the proof of Lemma 4.2. $\square$

4.2.3 Refinement for Small Items

Proof of Lemma 4.23: Define $r(i)=\frac{v_d(i)}{v_{\hat{d}}(i)}$ for any $i\in I$. Assume, without loss of generality, that $I\setminus L=\{1,2,\ldots,s\}$ for some $s\in\mathbb{N}$, and $r(1)\leq r(2)\leq\ldots\leq r(s)$.

If $\bar{a}\cdot(\bar{v}^1+\bar{v}^2)\leq\frac{1}{q^2}\mathrm{OPT}+2q\delta$ define $H_1=I\setminus L$ and $H_j=\emptyset$ for $j\in\{2,\ldots,q\}$. Let $Q\subseteq I\setminus L$ and $\beta\in[\frac{1}{q},1]$ which satisfies (71). We can select $X=I\setminus L$. It follows that $\|\mathbb{1}_{Q\setminus X}\bullet\bar{a}\bullet\bar{v}^d\|=0\leq\beta\cdot\bar{a}\cdot\bar{v}^d$ and $\|\mathbb{1}_X\bullet\bar{a}\bullet(\bar{v}^1+\bar{v}^2)\|=\bar{a}\cdot\left(\bar{v}^1+\bar{v}^2\right)\leq\frac{16}{q}\mathrm{OPT}+2q\delta$. This shows the statement of the lemma in case $\bar{a}\cdot(\bar{v}^1+\bar{v}^2)\leq\frac{1}{q^2}\mathrm{OPT}+2q\delta$. We henceforth assume that

$$\bar{a}\cdot(\bar{v}^1+\bar{v}^2)>\frac{1}{q^2}\mathrm{OPT}+2q\delta\ . \tag{87}$$

Define $h_0=0$, and for $j=1,\ldots,q$ set

$$h_j=\min\left\{i\in[s]\ \middle|\ \left(\bar{a}\wedge\mathbb{1}_{[i]}\right)\cdot(\bar{v}^1+\bar{v}^2)\geq\frac{j}{q}\cdot\bar{a}\cdot(\bar{v}^1+\bar{v}^2)\right\}\ . \tag{88}$$

Observe that the set over which the minimum is taken is non-empty for all $j\in\{1,\ldots,q\}$. Hence, h_j is well defined. Define $H_j=\{i\in\{1,\ldots,s\}\mid h_{j-1}<i\leq h_j\}$; then $H_j=\{1,\ldots,h_j\}\setminus\{1,\ldots,h_{j-1}\}$ for $j=1,\ldots,q$.

Let $Q\subseteq I\setminus L$ and $\beta\in[\frac{1}{q},1]$ satisfy (71). For $j=1,\ldots,q$ and $C\in\mathcal{C}$ it holds that $v_d(C\cap H_j)\leq 1$, and

$$v_d(C\cap H_j)=\sum_{i\in C\cap H_j}v_d(i)=\sum_{i\in C\cap H_j}v_{\hat{d}}(i)\cdot r(i)\leq r(h_j)\sum_{i\in C\cap H_j}v_{\hat{d}}(i)\leq r(h_j)\ .$$

Thus, $v_d(C\cap H_j)\leq\min\{1,r(h_j)\}$. We conclude that

$$\max\left\{v_d(C\cap H_j)\mid C\in\mathcal{C}\right\}\leq\min\{1,r(h_j)\} \tag{89}$$

for $j=1,\ldots,q$.

We use in our proof the following inequality (that we prove later), for $j=2,\ldots,q$:

$$\|\mathbb{1}_{H_j}\bullet\bar{a}\bullet\bar{v}^d\|\geq\frac{1}{2}\min\{1,r(h_{j-1})\}\cdot\frac{1}{q^3}\mathrm{OPT}, \tag{90}$$

For $j=1,\ldots,q$ define

$$\beta_j=\max\left\{0,\ \|\mathbb{1}_{Q\cap H_j}\bullet\bar{a}\bullet\bar{v}^d\|-\beta\|\mathbb{1}_{H_j}\bullet\bar{a}\bullet\bar{v}^d\|\right\}\ .$$

It follows from (71) and (89) that

$$\beta_j \leq \frac{\mathrm{OPT}}{q^5} \cdot \max\left\{v_d(C \cap H_j) \mid C \in \mathcal{C}\right\} \leq \min\{r(h_j), 1\} \cdot \frac{\mathrm{OPT}}{q^5} \ .$$

For every $j \in [q] \setminus \{1\}$ we define a set $X_j \subseteq Q \cap H_j$. If $\|\mathbb{1}_{Q\cap H_j} \bullet \bar{a} \bullet \bar{v}^d\| + \beta_{j-1} - \beta_j \leq \beta \cdot \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\|$ then we define $X_j = \emptyset$. Otherwise, we define X_j to be an inclusion-minimal subset of $Q \cap H_j$ such that $\|\mathbb{1}_{Q\cap H_j \setminus X_j} \bullet \bar{a} \bullet \bar{v}^d\| + \beta_{j-1} - \beta_j \leq \beta \cdot \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\|$. Observe that

$$\|\mathbb{1}_{Q\cap H_j \setminus (Q \cap H_j)} \bullet \bar{a} \bullet \bar{v}^d\| + \beta_{j-1} - \beta_j \leq \beta_{j-1} \leq \min\{1, \tau_{j-1}\} \cdot \frac{\mathrm{OPT}}{q^5} \leq \beta \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\|,$$

where the last inequality follows from $\beta \geq \frac{1}{q}$ and (90). Hence, there exists $X_j \neq \emptyset$. As the set is inclusion-minimal, it follows that there is $x_j \in X_j$ such that $\|\mathbb{1}_{X_j \setminus \{x_j\}} \bullet \bar{a} \bullet \bar{v}^d\| \leq \beta_{j-1} \leq \frac{\mathrm{OPT}}{q^5}$. Thus,

$$\begin{aligned}
\|\mathbb{1}_{X_j \setminus \{x_j\}} \bullet \bar{a} \bullet \bar{v}^{\hat{d}}\| &= \sum_{i \in X_j \setminus \{x_j\}} \bar{a}_i \cdot v_{\hat{d}}(i) = \sum_{i \in X_j \setminus \{x_j\}} \bar{a}_i \cdot \frac{v_d(i)}{r(i)} \leq \sum_{i \in X_j \setminus \{x_j\}} \bar{a}_i \cdot \frac{v_d(i)}{r(h_{j-1})} \\
&= \frac{\|\mathbb{1}_{X_j \setminus \{x_j\}} \bullet \bar{a} \bullet \bar{v}^d\|}{r(h_{j-1})} \leq \frac{\beta_{j-1}}{r(h_{j-1})} \leq \frac{1}{r(h_{j-1})} \min\{r(h_{j-1}), 1\} \cdot \frac{\mathrm{OPT}(I, v)}{q^5} \leq \frac{\mathrm{OPT}(I, v)}{q^5},
\end{aligned}$$

where the first inequality holds as $X_j \subseteq H_j$.

Define $X = (H_q \cap Q) \cup \bigcup_{j=2}^{q} X_j$. It follows that

$$\begin{aligned}
\|\mathbb{1}_{Q \setminus X} \cdot \bar{a} \cdot \bar{v}^d\| &= \sum_{j=1}^{q-1} \|\mathbb{1}_{(Q \setminus X) \cap H_j} \cdot \bar{a} \cdot \bar{v}^d\| \\
&= \|\mathbb{1}_{(Q \setminus X) \cap H_1} \cdot \bar{a} \cdot \bar{v}^d\| - \beta_1 + \sum_{j=2}^{q-1} \Big(\|\mathbb{1}_{(Q \setminus X) \cap H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \beta_{j-1} - \beta_j \Big) + \beta_{q-1} \\
&\leq \beta \sum_{j=1}^{q-1} \|\mathbb{1}_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \beta_{q-1} \\
&\leq \beta \sum_{j=1}^{q-1} \|\mathbb{1}_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \min\{r(h_{q-1}), 1\} \cdot \frac{\mathrm{OPT}(I, v)}{q^5} \\
&\leq \beta \sum_{j=1}^{q} \|\mathbb{1}_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| = \beta \cdot \bar{a} \cdot \bar{v}^d \ .
\end{aligned}$$

The first equality holds as $\mathrm{supp}(\bar{a}) \subseteq \bigcup_{j \in [q]} H_j$. The first inequality follows from the definitions of β_1 and X_j (for $j \in \{2, \ldots, q-1\}$). The last inequality follows from $\beta \geq \frac{1}{q}$ and (90).

Note that $\|\mathbb{1}_{H_q} \cdot \bar{a} \cdot \left(\bar{v}^1 + \bar{v}^2\right)\| \leq \frac{\bar{a} \cdot \bar{v}^d}{q} \leq \frac{2 \cdot \mathrm{OPT}}{q}$. Thus,

$$\begin{aligned}
\|\mathbb{1}_X \bullet \mathbb{1}_A \bullet (\bar{v}^1 + \bar{v}^2)\| &\leq \|\mathbb{1}_{H_q} \cdot \bar{a} \cdot \left(\bar{v}^1 + \bar{v}^2\right)\| + \sum_{j=2}^{q} \|\mathbb{1}_{X_j} \cdot \bar{a} \cdot \left(\bar{v}^1 + \bar{v}^2\right)\| \\
&\leq \frac{2 \cdot \mathrm{OPT}}{q} + q \cdot 2 \cdot \frac{\mathrm{OPT}}{q^5} + 2\delta q \leq \frac{16}{q}\mathrm{OPT} + 2\delta q \ .
\end{aligned}$$

It remains to show that (90) holds. For $j = 1, \dots, q$, we have

$$\begin{aligned}
\|\mathbb{1}_{H_j} \bullet \bar{a} \bullet (\bar{v}^1 + \bar{v}^2)\| &= \|\mathbb{1}_{[h_j]} \bullet \bar{a} \bullet (\bar{v}^1 + \bar{v}^2)\| - \|\mathbb{1}_{h_{j-1}} \bullet \bar{a} \bullet (\bar{v}^1 + \bar{v}^2)\| \\
&\geq \frac{j}{q}\bar{a} \cdot (\bar{v}^1 + \bar{v}^2) - \frac{j-1}{q}\bar{a} \cdot (\bar{v}^1 + \bar{v}^2) - 2\delta \\
&= \frac{1}{q}\bar{a} \cdot (\bar{v}^1 + \bar{v}^2) - 2\delta \\
&\geq \frac{1}{q}\left(\frac{1}{q^2}\mathrm{OPT} + 2\delta q\right) - 2\delta \\
&= \frac{1}{q^3}\mathrm{OPT}(I, v) \ .
\end{aligned} \tag{91}$$

The first inquality follows from (88) and $v_1(i) + v_2(i) \leq 2\delta$ for all $i \in I \setminus L$. The second inequality follows from (87). Additionally, for $j = 2, \dots, \ell$ we have

$$\begin{aligned}
\|\mathbb{1}_{H_j} \bullet \bar{a} \bullet (\bar{v}^1 + \bar{v}^2)\| &= \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\| + \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^{\hat{d}}\| \\
&= \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\| + \sum_{i \in H_j} \bar{a}_i \cdot v_{\hat{d}}(i) \\
&= \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\| + \sum_{i \in H_j} \bar{a}_i \cdot \frac{v_d(i)}{r(i)} \\
&\leq \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\| + \sum_{i \in H_j} \bar{a}_i \cdot \frac{v_d(i)}{r(h_{j-1})} \\
&= \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\| \cdot \left(1 + \frac{1}{r(h_{j-1})}\right),
\end{aligned} \tag{92}$$

where the inequality follows from $r(1) \leq r(2) \leq \ldots \leq r(p)$. Using (91) and (92), we get

$$\forall j = 2, \dots, q : \qquad \|\mathbb{1}_{H_j} \bullet \bar{a} \bullet \bar{v}^d\| \geq \left(1 + \frac{1}{r(h_{j-1})}\right)^{-1} \cdot \frac{1}{q^3}\mathrm{OPT} \geq \frac{1}{2}\min\{1, \tau_{j-1}\} \cdot \frac{1}{q^3}\mathrm{OPT},$$

where the inequality follows from $\left(1 + x^{-1}\right)^{-1} \geq \frac{1}{2}\min\{1, x\}$ for every $x \geq 0$. Inequality (90) follows from the last inequality. $\square$

4.3 Existence of ψ-Relaxations

In this section we prove Lemmas 4.4 to 4.6. That is, we show how to obtain relaxations for various configurations.

Proof of Lemma 4.4: Let $S \subseteq C \setminus L$ be an inclusion-minimal set such that either $v_1(C \setminus S) \leq 1 - \delta$ or $v_2(C \setminus S) \leq 1 - \delta$. As S is inclusion-minimal, it holds that

$$\forall i \in S : \qquad v\Big(C \setminus (S \setminus \{i\})\Big) > (1 - \delta, 1 - \delta). \tag{93}$$

Such a set exists, since $C \in \mathcal{C}_0$.

In the following we show that $v(S) \leq (2\delta, 2\delta)$. Suppose, for sake of contradiction, that $v_1(S) > 2\delta$ or $v_2(S) > 2\delta$. Then $S \neq \emptyset$ and there is an $i \in S$. Assume, without loss of generality, that $v_1(S) > 2\delta$. Then $v_1(S \setminus \{i\}) > \delta$ as all items in S are small, and $i \in S$. Therefore,

$$v_1\Big(C \setminus (S \setminus \{i\})\Big) = v_1(C) - v_1(S \setminus \{i\}) \leq 1 - \delta,$$

contradicting (93). Thus, $v(S) \leq (2\delta, 2\delta)$.

Define $C_1 = C \setminus S$ and $C_2 \in \mathcal{C}^*$ by

$$C_2(i) = \begin{cases} \kappa, & i \in S, \\ 0, & i \notin S \end{cases}$$

for $i \in I$, where $\kappa = \left\lfloor \frac{1}{2}(\delta^{-1} - 1) \right\rfloor$. Observe that C_1 has δ-slack by definition of S. Additionally,

$$v_1(C_2) \leq v_1(S) \cdot \kappa \leq 2\delta\kappa \leq 2\delta \cdot \frac{1}{2}(\delta^{-1} - 1) \leq 1 - \delta,$$

thus C_2 is a multi-configuration with δ-slack.

Define $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ by $\bar{\lambda}_{C_1} = 1$, $\bar{\lambda}_{C_2} = \frac{1}{\kappa}$ and $\bar{\lambda}_{C'} = 0$ for $C' \in \mathcal{C} \setminus \{C_1, C_2\}$. Clearly, for any $C' \in \mathcal{C}^*$ such that $\bar{\lambda}_{C'} > 0$ it holds that C' has δ-slack. Thus, $\bar{\lambda}$ has δ-slack.

For any $i \in C \setminus S$ we have

$$\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = C_1(i) + \frac{1}{\kappa} \cdot C_2(i) = 1 + 0 = 1 \ .$$

For any $i \in S$ it holds that

$$\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = C_1(i) + \frac{1}{\kappa} \cdot C_2(i) = 0 + \frac{1}{\kappa} \cdot \kappa = 1 \ .$$

For any $i \in I \setminus C$ it holds that

$$\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = C_1(i) + \frac{1}{\kappa} \cdot C_2(i) = 0 + \frac{1}{\kappa} \cdot 0 = 0 \ .$$

Since $\delta^{-1} \in \mathbb{N}$, we have $\kappa \geq \frac{1}{2}(\delta^{-1} - 1) - \frac{1}{2} = \frac{1}{2}\delta^{-1} - 1$. Therefore,

$$\|\bar{\lambda}\| = \sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} = \bar{\lambda}_{C_1} + \bar{\lambda}_{C_2} = 1 + \frac{1}{\kappa} \leq 1 + \frac{1}{\frac{1}{2}\delta^{-1} - 1} = 1 + \frac{2\delta}{1 - 2\delta} \leq 1 + 4\delta,$$

where the last inequality holds as $\delta \leq 0.1$

We showed that $\bar{\lambda}$ is a $(1 + 4\delta)$-relaxation of C. This completes the proof of the lemma. ☐

Proof of Lemma 4.5: Let $C \cap L = \{i_1, \ldots, i_h\}$. Define h configurations $C_1, \ldots, C_h$ by $C_\ell = C \setminus \{i_\ell\}$ for $\ell = 1, \ldots, h-1$ and $C_h = C \cap L \setminus \{i_h\}$. It can be easily shown that $C_1, \ldots, C_h$ are configurations. Define $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ by

$$\bar{\lambda}_{C'} = \begin{cases} \frac{1}{h-1}, & C' = C_\ell \text{ for some } h \in \{1, \ldots, \ell\}, \\ 0, & \text{otherwise} \ . \end{cases}$$

For $\ell = 1, \ldots, h$ it holds that i_ℓ is large; thus, there is $d_\ell \in \{1, 2\}$ such that $v_{d_\ell}(i_\ell) \geq \delta$. Therefore,

$$v_{d_\ell}(C_\ell) \leq v_{d_\ell}(C \setminus \{i_\ell\}) = v_{d_\ell}(C) - v_{d_\ell}(i_\ell) \leq 1 - \delta \ .$$

That is, all configurations $C_1, \ldots, C_h$ have δ-slack. Thus, for any $C' \in \mathcal{C}^*$ with $\bar{\lambda}_{C'} > 0$ it holds that C' has δ-slack. Hence, $\bar{\lambda}$ has δ-slack.

For any $i \in C \cap L$ there is an $\ell \in \{1, \ldots, h\}$ such that $i = i_\ell$. Thus,

$$\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = \sum_{j=1}^{h} \frac{1}{h-1} \cdot C_j(i_\ell) = \sum_{j \in [h] \setminus \{\ell\}} \frac{1}{h-1} = 1 \ .$$

For any $i \in C \setminus L$ it holds that $i \in C_\ell$ for $\ell = 1, \ldots, h-1$; thus,

$$\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = \sum_{j=1}^{h} \frac{1}{h-1} \cdot C_j(i) = \sum_{j=1}^{h-1} \frac{1}{h-1} = 1 \ .$$

For any $i \in I \setminus C$ we have $i \notin C_\ell$ for $\ell = 1, \ldots, h$. Therefore,

$$\sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} \cdot C'(i) = \sum_{j=1}^{h} \frac{1}{h-1} \cdot C_j(i) = 0 \ .$$

Finally,

$$\|\bar{\lambda}\| = \sum_{C' \in \mathcal{C}^*} \bar{\lambda}_{C'} = \sum_{\ell=1}^{h} \bar{\lambda}_{C_\ell} = \frac{h}{h-1} \ .$$

Thus, we showed that $\bar{\lambda}$ is a $\frac{h}{h-1}$-relaxation of C. □

Proof of Lemma 4.6: Define $C' \in \mathcal{C}^*$ by

$$C'(i) = \begin{cases} \kappa, & i \in C, \\ 0, & \text{otherwise,} \end{cases}$$

where $\kappa = \lceil \frac{1}{2}\delta^{-1} \rceil$ and $\bar{\lambda} \in [0,1]^{\mathcal{C}^*}$ by $\bar{\lambda}_{C'} = \frac{1}{\kappa}$ and $\bar{\lambda}_D = 0$ for any $D \in \mathcal{C}^* \setminus \{C'\}$. Observe that

$$v_1(C') = \sum_{i \in I} v_1(i) \cdot C'(i) = \kappa \cdot v_1(C) \leq \left\lceil \frac{1}{2}\delta^{-1} \right\rceil \cdot \delta \leq \left(\frac{1}{2} \cdot \delta^{-1} + 1 \right) \cdot \delta \leq \frac{1}{2} + \delta \leq 0.6 \leq 1 - \delta,$$

where the last two inequalities follow from $\delta \in (0, 0.1)$. Thus, C' has δ-slack and hence $\bar{\lambda}$ is with δ-slack.

For any $i \in C$ it holds that $\sum_{D \in \mathcal{C}^*} \bar{\lambda}_D \cdot D(i) = \bar{\lambda}_{C'} \cdot C'(i) = \frac{1}{\kappa} \cdot \kappa = 1$. Also, for any $i \in I \setminus C$ it holds that $\sum_{D \in \mathcal{C}^*} \bar{\lambda}_D \cdot D(i) = \bar{\lambda}_{C'} \cdot C'(i) = 0$. Finally,

$$\|\bar{\lambda}\| = \frac{1}{\kappa} \leq \frac{1}{\lceil \frac{1}{2}\delta \rceil} \leq 2\delta \leq 4\delta \ .$$

Thus, $\bar{\lambda}$ is a 4δ-relaxation of C, as required. □

4.4 Solving the Matching-LP

In this section we present a PTAS for the MLP problem, thus proving Lemma 1.8. Let $\delta \in (0, 0.1)$ and $\varepsilon \in (0, 0.1)$. Our objective is to obtain a polynomial-time $(1 + O(\varepsilon))$-approximation for MLP. To this end we use a result of Grötschel, Lovász, and Schrijver [GLS81], which outlines the ellipsoid method via separation oracles. A separation oracle for a polytope $P \subseteq \mathbb{R}^n$ accepts as input a point $\bar{x} \in \mathbb{R}^n$, and either determines that $\bar{x} \in P$ or finds $\bar{c} \in \mathbb{R}^n$ such that $\bar{x} \cdot \bar{c} < \bar{y} \cdot \bar{c}$ for any $\bar{y} \in P$. That is, the oracle finds a hyperplane which separates between $\bar{x}$ and the polytope P. It is also required that the encoding size of the returned hyperplane is polynomial in the query encoding size. Given a separation oracle, the ellipsoid method either determines that $P = \emptyset$ or finds $\bar{x} \in P$ in time polynomial in n and the *facet complexity* of P. As a consequence, if $P = \emptyset$ then the execution of the ellipsoid method is comprised of invocations of the separation oracle that always result in a separating hyperplane. If $P \neq \emptyset$, then at least one of the calls to the separation oracle results in $\bar{x} \in P$.

We use an approximate variant of the separation oracle commonly used to solve linear programs similar to (1) (see, e.g., [KK82]). In the classic setting, the ellipsoid method is executed with the dual of the original linear program, as this program has a polynomial number of variables. For example, the dual linear program of (1) has $|I|$ variables. This approach cannot be directly implemented for MLP, since the number of variables in both the primal and dual linear programs is non-polynomial in the δ-huge free 2VBP instance (I, v), due to the number of linear constraints required to represent the matching polytop. We overcome this difficulty by projecting polytopes in a vector space of non-polynomial dimension into polytopes with polynomial dimension. A similar approach was recently used by Fairstein et al. [FKS21].

We use the following definitions and lemmas from Grötschel et al. [GLS88].

Definition 4.31 ([GLS88, Definition 6.2.2]). *Let $P \subseteq \mathbb{R}^n$ be a polyhedron, and $\varphi \geq n+1$ a positive integer.*

1. *We say that P has* facet complexity at most φ *if there exists a system of linear inequalities with rational coefficients that has a solution set P that the encoding length of each inequality in the system is at most φ.*
2. *We say that P has a* vertex complexity at most φ *if there exist finite sets V_1, V_2 of rational vectors such that $P = \mathtt{conv}(V_1) + \mathtt{cone}(V_2)$ and each of the vectors in $V_1 \cup V_2$ has encoding length at most φ.*[13]
3. *A* well-described polyhedron *is a triplet (P, n, φ) where $P \subseteq \mathbb{R}^n$ is a polyhedron with facet complexity at most φ.*

Lemma 4.32 ([GLS88, Lemma 6.2.4]). *Let $P \subseteq \mathbb{R}^n$ be a polyhedron with facet complexity at most φ. Then P has vertex complexity at most $4n^2 \cdot \varphi$.*

Proposition 4.33 (The *Ellipsoid Method*, [GLS88, Theorem 6.4.1]). *There is an algorithm* Ellipsoid *which given n, φ and a separation oracle for a well-described polyhedron (P, n, φ), determines that either $P = \emptyset$ or returns $\bar{x} \in P$ in time polynomial in $n + \varphi$.*

Throughout this section, we define multiple mathematical optimization problems. We use $\text{OPT}(\mathcal{P})$ to denote the value of the optimal solution for the problem $\mathcal{P}$. We use $\langle x \rangle$ to denote the encoding length of a number/vector/inequality x. To simplify notation, we assume the δ-2VBP instance (I, v) is fixed throughout this section, and omit it from the input of the algorithms. We use $G = (L, E)$ to denote the δ-matching graph of (I, v) as defined in Section 1.3, and $P_{\mathcal{M}}(G)$ is the matching polytope of G. Recall that $\mathcal{E}$ is the projection function defined in Section 1.3.

We first simplify our problem. We relax the requirement $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = 1$ in (4) and use

[13] $\mathtt{conv}(V_1)$ is the *convex hull* of V_1 and $\mathtt{cone}(V_2)$ is the *conic hull* of V_2. We refer the reader to Grötschel et al. [GLS88] for the formal definitions.

inequality instead. That is,

$$\begin{aligned}
\text{rMLP}: \quad & \min && \sum_{C\in\mathcal{C}} \bar{x}_C \\
& \forall i\in I: && \sum_{C\in\mathcal{C}} \bar{x}_C \cdot C(i) \geq 1 \\
& && \mathcal{E}(\bar{x}) \in P_{\mathcal{M}}(G) \\
& \forall C\in\mathcal{C}: && \bar{x}_C \geq 0
\end{aligned} \tag{94}$$

It can be easily shown that the optima of (4) and (94) are equal; furthermore, a solution for (94) can be easily converted to a solution for (4) of the same or lower value.

Our objective is to find a variant of (94) in which the set $\mathcal{C}$ is replaced by a polynomial-size set $\mathcal{D} \subseteq \mathcal{C}$, while approximately preserving the optimal value. To this end we use the following family of polytopes:

$$\forall \mathcal{D}\subseteq\mathcal{C}: \quad P(\mathcal{D}) = \left\{ (\bar{x},\bar{y}) \;\middle|\; \begin{array}{c} \bar{x}\in\mathbb{R}^{\mathcal{D}}_{\geq 0},\ \bar{y}\in P_{\mathcal{M}}(G) \\ \mathcal{E}(\bar{x}) \leq \bar{y} \\ \forall i\in I: \ \sum_{C\in\mathcal{D}} \bar{x}_C \cdot C(i) \geq 1 \end{array} \right\} . \tag{95}$$

Given $\mathcal{D} \subseteq \mathcal{C}$, with a slight abuse of notation we refer to a vector $\bar{x} \in \mathbb{R}^{\mathcal{D}}_{\geq 0}$ as a vector in $\mathbb{R}^{\mathcal{C}}_{\geq 0}$ where $\bar{x}_C = 0$ for every $C \in \mathcal{C} \setminus \mathcal{D}$. This ensures that the term $\mathcal{E}(\bar{x})$ is well defined. Since $P_{\mathcal{M}}(G)$ is downward closed, we have that rMLP is equivalent to the problem of finding $(\bar{x}, \bar{y}) \in P(\mathcal{C})$ such that $\|\bar{x}\|$ is minimized.[14] For $\mathcal{D} \subseteq \mathcal{C}$ we define rMLP($\mathcal{D}$) as the problem of finding $(\bar{x}, \bar{y}) \in P(\mathcal{D})$ such that $\|\bar{x}\|$ is minimized. It follows that OPT(rMLP($\mathcal{D}$)) $\geq$ OPT(rMLP) for any $\mathcal{D} \subseteq \mathcal{C}$.

We use $P(\mathcal{D})$ to define a family of additional polytopes $Q(\mathcal{D}, h)$ in $\mathbb{R}^E$, one for each $\mathcal{D} \subseteq \mathcal{C}$ and $h \in \mathbb{R}_{\geq 0}$:

$$Q(\mathcal{D},h) = \left\{ \bar{y}\in\mathbb{R}^E \;\middle|\; \exists \bar{x}\in\mathbb{R}^{\mathcal{D}}_{\geq 0}: \ (\bar{x},\bar{y})\in P(\mathcal{D}) \text{ and } \|\bar{x}\| \leq h \right\} . \tag{96}$$

It thus follows that $Q(\mathcal{D}, h) \neq \emptyset$ if and only if OPT(rMLP($\mathcal{D}$)) $\leq h$. Furthermore, $Q(\mathcal{D}, h)$ is a polytope in a vector space of polynomial size. We use the ellipsoid method to determine if $Q(\mathcal{C}, h) = \emptyset$ for various values of h. The separation oracle first checks if $\bar{y} \in P_{\mathcal{M}}(G)$, and otherwise finds a separating hyperplane using a separation oracle for the matching polytope. If $\bar{y} \in P_{\mathcal{M}}(G)$ we use the following linear program, which depends on $\bar{y} \in P_{\mathcal{M}}(G)$ and $\mathcal{D} \subseteq \mathcal{C}$, to obtain a separating hyperplane:

$$\begin{aligned}
\text{PRIMAL}(\bar{y},\mathcal{D}) \quad & \min && \sum_{C\in\mathcal{D}} \bar{x}_C, \\
& \forall i\in I: && \sum_{C\in\mathcal{D}} \bar{x}_C \cdot C(i) \geq 1, \\
& \forall e\in E: && \sum_{C\in S(e)\cap\mathcal{D}} \bar{x}_C \leq \bar{y}_e, \\
& \forall C\in\mathcal{C}: && \bar{x}_C \geq 0 \ .
\end{aligned} \tag{97}$$

where for every $e \in E$ we define its superset of configurations as $S(e) = \{C \in \mathcal{C} \mid e \subseteq C\}$. Using this notation it holds that $(\mathcal{E}(\bar{x}))_e = \sum_{C\in S(e)} \bar{x}_C$. It follows that $\bar{y} \in Q(\mathcal{D}, h)$ if and only if $\bar{y} \in P_{\mathcal{M}}(G)$ and OPT(PRIMAL($\bar{y}, \mathcal{D}$)) $\leq h$.

[14] A polytope $P \subseteq \mathbb{R}^n_{\geq 0}$ is *downward closed* if for any $\bar{x} \in P$ and $\bar{y} \in \mathbb{R}^n_{\geq 0}$ such that $\bar{y} \leq \bar{x}$ it holds that $\bar{y} \in P$.

Recall the set $\mathcal{C}_2$ is defined in (2). For any $C \in \mathcal{C}$ it holds that $C \in \mathcal{C}_2$ if and only if there is $e \in E$ such that $C \in S(e)$. We use this observation to derive the dual of PRIMAL$(\bar{y}, \mathcal{D})$, which is the following linear program:

$$
\begin{array}{llll}
\mathrm{DUAL}(\bar{y},\mathcal{D}) & \max & \sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e, & \\
& \forall C \in \mathcal{D} \setminus \mathcal{C}_2 : & \sum_{i\in C} \bar{\lambda}_i \leq 1, & \\
& \forall e \in E,\ C \in S(e) \cap \mathcal{D} : & \sum_{i\in C} \bar{\lambda}_i \leq 1 + \beta_e, & (98)\\
& \forall i \in I : & \bar{\lambda}_i \geq 0, & \\
& \forall e \in E : & \bar{\beta}_e \geq 0 \ . &
\end{array}
$$

Observe that the feasibility region of DUAL$(\bar{y}, \mathcal{D})$ is independent of $\bar{y}$. That is, for any $\mathcal{D} \subseteq \mathcal{C}$ we can define

$$
R(\mathcal{D}) = \left\{ (\bar{\lambda}, \bar{\beta}) \in \mathbb{R}^I_{\geq 0} \times \mathbb{R}^E_{\geq 0} \;\middle|\; \begin{array}{ll} \forall C \in \mathcal{D} \setminus \mathcal{C}_2 : & \sum_{i\in C} \bar{\lambda}_i \leq 1 \\ \forall e \in E,\ C \in S(e) \cap \mathcal{D} : & \sum_{i\in C} \bar{\lambda}_i \leq 1 + \beta_e \end{array} \right\} . \tag{99}
$$

Then DUAL$(\bar{y}, \mathcal{D})$ is the problem of finding $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{D})$ for which $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e$ is maximized.

We use the following relation between $R(\mathcal{C})$ and $Q(\mathcal{C}, h)$ to generate separating hyperplanes.

Lemma 4.34. *For any $h \in \mathbb{R}_{\geq 0}$, $\bar{y} \in Q(\mathcal{C}, h)$ and $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$ it holds that*

$$
\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq h \ .
$$

Proof. As $\bar{y} \in Q(\mathcal{C}, h)$ it follows that OPT(DUAL$(\bar{y}, \mathcal{C})$) = OPT(PRIMAL$(\bar{y}, \mathcal{C})$) $\leq h$. Thus, as $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$ we have

$$
\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq \mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{C})) \leq h \ . \qquad \square
$$

We also use $R(\mathcal{C})$ to bound the facet complexity of $Q(\mathcal{C}, h)$.

Lemma 4.35. *There is a polynomial p_1 (independent of the instance (I, v)) such that for any $h \geq 0$ the facet complexity of $Q(\mathcal{C}, h)$ is at most $p_1(|I| + \langle h \rangle)$.*

Proof. By (99), the facet complexity of $R(\mathcal{C})$ is polynomial in the encoding of the input instance (I, v). Therefore, by Lemma 4.32, the vertex complexity of $R(\mathcal{C})$ is at most $4 \cdot (|I| + |I|^2)$ times the facet complexity of $R(\mathcal{C})$. Thus, the vertex complexity of $R(\mathcal{C})$ is polynomial in $|I|$. Hence, there is a polynomial q such that the vertex complexity of $R(\mathcal{C})$ is at most $q(|I|)$.

By Definition 4.31 there are $V_1, V_2 \subseteq \mathbb{R}^I_{\geq 0} \times \mathbb{R}^E_{\geq 0}$ such that $R(\mathcal{C}) = \mathtt{conv}(V_1) + \mathtt{cone}(V_2)$ and $\langle \bar{u} \rangle \leq q(|I|)$ for every $\bar{u} \in V_1 \cup V_2$. For any $h \geq 0$ define

$$
Q'(h) = \left\{ \bar{y} \in P_{\mathcal{M}}(G) \;\middle|\; \begin{array}{ll} \forall (\bar{\lambda}, \bar{\beta}) \in V_1 : & \sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq h \\ \forall (\bar{\lambda}, \bar{\beta}) \in V_2 : & \sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq 0 \end{array} \right\}
$$

Claim 4.36. *For any $h \geq 0$ it holds that $Q(\mathcal{C}, h) \subseteq Q'(h)$.*

Proof. Let $\bar{y} \in Q(\mathcal{C}, h)$. For any $(\bar{\lambda}, \bar{\beta}) \in V_1$ it holds that $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$, thus $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq h$ by Lemma 4.34. Suppose, for sake of contradiction, that there is $(\bar{\lambda}, \bar{\beta}) \in V_2$ such that $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e = \xi > 0$. It therefore holds that $(\frac{h+1}{\xi}\bar{\lambda}, \frac{h+1}{\xi}\bar{\beta}) \in R(\mathcal{C})$. Thus

$$h \geq \mathrm{OPT}(\mathrm{PRIMAL}(\bar{y}, \mathcal{C})) = \mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{C})) \geq \sum_{i\in I} \frac{h+1}{\xi} \cdot \bar{\lambda}_i - \sum_{e\in E} \frac{h+1}{\xi} \cdot \bar{\beta}_e \cdot \bar{y}_e \geq h+1,$$

a contradiction. Hence, $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq 0$ for every $(\bar{\lambda}, \bar{\beta}) \in V_2$, and $\bar{y} \in Q'(h)$. $\diamond$

Claim 4.37. *For any $h \geq 0$ it holds that $Q'(h) \subseteq Q(\mathcal{C}, h)$.*

Proof. Let $\bar{y} \in Q'(h)$ and $(\bar{\lambda}^*, \bar{\beta}^*) \in R(\mathcal{C})$. As $R(\mathcal{C}) = \mathtt{conv}(V_1) + \mathtt{cone}(V_2)$ there are numbers $\zeta_{\bar{\lambda},\bar{\beta}} \geq 0$ for all $(\bar{\lambda}, \bar{\beta}) \in V_1$ and $\xi_{\bar{\lambda},\bar{\beta}} \geq 0$ for all $(\bar{\lambda}, \bar{\beta}) \in V_2$ such that $\sum_{(\bar{\lambda},\bar{\beta})\in V_1} \zeta_{\bar{\lambda},\bar{\beta}} = 1$, and

$$(\bar{\lambda}^*, \bar{\beta}^*) = \sum_{(\bar{\lambda},\bar{\beta})\in V_1} \zeta_{\bar{\lambda},\bar{\beta}} \cdot (\bar{\lambda}, \bar{\beta}) + \sum_{(\bar{\lambda},\bar{\beta})\in V_2} \xi_{\bar{\lambda},\bar{\beta}} \cdot (\bar{\lambda}, \bar{\beta}) \ .$$

Thus,

$$\begin{aligned}
& \sum_{i\in I} \bar{\lambda}^*_i - \sum_{e\in E} \bar{\beta}^*_e \cdot \bar{y}_e \\
= \ & \sum_{(\bar{\lambda},\bar{\beta})\in V_1} \zeta_{\bar{\lambda},\bar{\beta}} \left(\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \right) + \sum_{(\bar{\lambda},\bar{\beta})\in V_2} \xi_{\bar{\lambda},\bar{\beta}} \left(\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \right) \\
\leq \ & \sum_{(\bar{\lambda},\bar{\beta})\in V_1} \zeta_{\bar{\lambda},\bar{\beta}} \cdot h + \sum_{(\bar{\lambda},\bar{\beta})\in V_2} \xi_{\bar{\lambda},\bar{\beta}} \cdot 0 \\
\leq \ & h \ .
\end{aligned}$$

That is, we showed that $\sum_{i\in I} \bar{\lambda}^*_i - \sum_{e\in E} \bar{\beta}^*_e \cdot \bar{y}_e \leq h$ for every $(\bar{\lambda}^*, \bar{\beta}^*) \in R(\mathcal{C})$. Hence, $\mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{C})) \leq h$. As it also holds that $\bar{y} \in P_{\mathcal{M}}(G)$, we conclude that $\bar{y} \in Q(\mathcal{C}, h)$. $\diamond$

By Claim 4.36 and Claim 4.37 it follows that $Q'(h) = Q(\mathcal{C}, h)$. Furthermore, by *Edmonds' matching polytope theorem* (see, e.g., Corollary 25.1a in Schrijver's book [Sch03]) it holds that

$$P_{\mathcal{M}}(G) = \left\{ \bar{y} \in \mathbb{R}^E_{\geq 0} \ \middle| \ \begin{array}{lll} \forall i \in L & : & \sum_{(i,i')\in E} \bar{x}_{(i,i')} \leq 1 \\ \forall U \subseteq L \text{ s.t. } U \text{ is odd} & : & \sum_{(i,i')\in E \text{ s.t. } i,i'\in U} \bar{x}_{(i,i')} \leq \left\lfloor \frac{|U|}{2} \right\rfloor \end{array} \right\}.$$

Thus,

$$Q(\mathcal{C}, h) = Q'(h) = \left\{ \bar{y} \in \mathbb{R}^E_{\geq 0} \ \middle| \ \begin{array}{lll} \forall (\bar{\lambda}, \bar{\beta}) \in V_1 & : & \sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq h \\ \forall (\bar{\lambda}, \bar{\beta}) \in V_2 & : & \sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \leq 0 \\ \forall i \in L & : & \sum_{(i,i')\in E} \bar{x}_{(i,i')} \leq 1 \\ \forall U \subseteq L \text{ s.t. } U \text{ is odd} & : & \sum_{(i,i')\in E \text{ s.t.} i,i'\in U} \bar{x}_{(i,i')} \leq \left\lfloor \frac{|U|}{2} \right\rfloor \end{array} \right\}.$$

That is, $Q(\mathcal{C}, h)$ is the solution set for a system of linear equations in which the encoding length of each inequality is at most $q(|I|) + \langle h \rangle + O(|I|^2)$. This completes the proof of Lemma 4.35. $\square$

Let $\mathcal{M}^*$ be a maximum matching in the graph G. Since each of the vertices in a matching polytope corresponds to a(n integral) matching, it holds that

$$\sum_{e \in E} \bar{y}_e \leq |\mathcal{M}^*| \quad \text{for all } \bar{y} \in P_{\mathcal{M}}(G) \ . \tag{100}$$

Since for every $e \in \mathcal{M}^*$ it holds that $e \in \mathcal{C}_2$, i.e., $v_1(e) > (1-\delta)$, for every solution $\bar{x}$ of rMLP we have

$$\begin{aligned} \sum_{C \in \mathcal{C}} \bar{x}_C &\geq \sum_{C \in \mathcal{C}} \bar{x}_C \cdot v_1(C) \geq \sum_{C \in \mathcal{C}} \bar{x}_C \sum_{i \in I} v_1(i) \cdot C(i) \\ &= \sum_{i \in I} v_1(i) \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \geq \sum_{i \in I} v_1(i) \geq \sum_{e \in \mathcal{M}^*} v_1(e) > (1-\delta)|\mathcal{M}^*| \ . \end{aligned}$$

Hence,

$$\mathrm{OPT}(\mathrm{rMLP}) > (1-\delta)|\mathcal{M}^*| \ .$$

We combine Lemma 4.34 with the next lemma that is proved later in this section.

Lemma 4.38. *There is a polynomial-time algorithm* Ellipsoid_R *which, given* $\bar{y} \in P_{\mathcal{M}}(G)$ *and* $h > (1-\delta)|\mathcal{M}^*|$*, returns*

- *either a subset* $\mathcal{D} \subseteq \mathcal{C}$ *of size* $|\mathcal{D}|$ *polynomial in the input size such that* $\mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{D})) \leq (1+\varepsilon)\, h$,
- *or a point* $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$ *such that* $\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e > h$.

Algorithm 3: Q_separator

Input : $\bar{y} \in \mathbb{R}^E_{\geq 0}$, $h > (1-\delta)|\mathcal{M}^*|$.
Output: Either a separating hyperplane between $Q(\mathcal{C}, h)$ and $\bar{y}$, or a subset $\mathcal{D} \subseteq \mathcal{C}$.

```
1 If ȳ ∉ P_M(G), then find a separating hyperplane between ȳ and P_M(G) and return it.
2 Run Ellipsoid_R (Lemma 4.38) with ȳ and h as its inputs
3 if Ellipsoid_R returned (λ̄, β̄) ∈ R(C) such that Σ_{i∈I} λ̄_i − Σ_{e∈E} β̄_e · ȳ_e > h then
4 |  return Σ_{i∈I} λ̄_i − Σ_{e∈E} β̄_e · z̄_e = h as a separating hyperplane
5 else
6 |  notify the ellipsoid algorithm to abort, and return the set D ⊆ C returned by Ellipsoid_R
7 end
```

We use algorithm Ellipsoid_R in Lemma 4.38 to derive a separation oracle for $Q(\mathcal{C}, h)$. The pseudocode of the oracle is given in Algorithm 3. We note there is a polynomial-time separation oracle for the matching polytope (see, e.g, Schrijver [Sch03]); thus, Step 3 can be implemented in polynomial time. While the algorithm does not formally qualify as a separation oracle, it gives the following guarantee:

Lemma 4.39. *Given* $\bar{y} \in \mathbb{R}^E_{\geq 0}$ *and* $h > (1-\delta)|\mathcal{M}^*|$*, Algorithm 3,*

- *either returns a separating hyperplane between* $Q(\mathcal{C}, h)$ *and* $\bar{y}$*,*
- *or notifies the ellipsoid method to abort and returns* $\mathcal{D} \subseteq \mathcal{C}$ *of polynomial cardinality such that* $\mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{D})) \leq (1+\varepsilon)h$*. In this case, it must hold that* $\bar{y} \in P_{\mathcal{M}}(G)$*.*

Proof. If $\bar{y} \notin P_{\mathcal{M}}(G)$ then Algorithm 3 finds a separating hyperplane between $\bar{y}$ and $P_{\mathcal{M}}(G)$. As $Q(\mathcal{C}, h) \subseteq P_{\mathcal{M}}(G)$, this hyperplane also separates between $\bar{y}$ and $Q(\mathcal{C}, h)$.

If the invocation of Ellipsoid_R returns $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$ such that $\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e > h$, then $\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{z}_e = h$ is a separating hyperplane between $\bar{y}$ and $Q(\mathcal{C}, h)$ by Lemma 4.34. Otherwise, by Lemma 4.38, the invocation of Ellipsoid_R returns a subset $\mathcal{D} \subseteq \mathcal{C}$ of polynomial cardinality such that $\text{OPT}(\text{DUAL}(\bar{y}, \mathcal{D})) \leq (1+\varepsilon)\, h$. It follows that in this case Algorithm 3 notifies the ellipsoid to abort and returns $\mathcal{D}$. ◻

Algorithm 4 utilizes Q_separator as a separation oracle. The algorithm may return a vector $\bar{x} \in \mathbb{R}^{\mathcal{D}}_{\geq 0}$ for some $\mathcal{D} \subseteq \mathcal{C}$. Recall that we interpret such a vector as a vector in $\mathbb{R}^{\mathcal{C}}$ as well.

Algorithm 4: Ellipsoid_Q

Input : $h > (1-\delta)|\mathcal{M}^*|$

Output: Either determine that $\text{OPT}(\text{rMLP}) > h$, or return a solution $\bar{x}'$ for rMLP with $\|\bar{x}'\| \leq (1+\varepsilon)h$.

1 Run Ellipsoid with $n = |E|$, $\varphi = p_1(|I| + \langle h \rangle)$ and Q_separator (and h) as the separation oracle
2 **if** *the ellipsoid method returned that the polytope is empty* **then**
3 | Return $\text{OPT}(\text{rMLP}) > h$
4 **else**
5 | This case can only happen if the Q_separator notified the ellipsoid to abort and returned a set $\mathcal{D} \subseteq \mathcal{C}$. Find an optimal solution $(\bar{x}', \bar{y}')$ for $\text{rMLP}(\mathcal{D})$ and return $\bar{x}'$.
6 **end**

Lemma 4.40. *In polynomial time, Algorithm 4 either determines that* $\text{OPT}(\text{rMLP}) > h$*, or finds a solution* $\bar{x}'$ *for* rMLP *satisfying* $\|\bar{x}'\| \leq (1+\varepsilon)h$.

Proof. By Lemma 4.35 it holds that (Q, n, φ) is a well-described polyhedron. As n, φ are polynomial in the instance, it follows the execution time of the ellipsoid method is polynomial. Furthermore, if the algorithm solves $\text{rMLP}(\mathcal{D})$ in Line 5 then, by Lemma 4.39, we have that $|\mathcal{D}|$ is polynomial, and hence $\text{rMLP}(\mathcal{D})$ can be solved in polynomial time (as there is a separation oracle for $\mathcal{E}(\bar{x}) \in P_{\mathcal{M}}(G)$, and the number of variables and additional constraints is polynomial).

By Lemma 4.39, if the ellipsoid method asserts that the polytope is empty, it holds that all invocations of Q_separator returned a separating hyperplane. Hence, this is a valid execution of Ellipsoid with a separation oracle for $Q(\mathcal{C}, h)$. It follows that $Q(\mathcal{C}, h) = \emptyset$, implying that $\text{OPT}(\text{rMLP}) = \text{OPT}(\text{rMLP}(\mathcal{C})) > h$ due to (96).

Otherwise, it must hold that the execution of the ellipsoid method was aborted by Q_separator at some iteration. Let $\bar{y} \in P_{\mathcal{M}}(G)$ be the value of $\bar{y}$ used in the call to Q_separator in this iteration, let $\mathcal{D} \subseteq \mathcal{C}$ be the subset of configurations returned by Q_separator, and let $(\bar{x}', \bar{y}') \in P(\mathcal{D})$ be the solution found in Line 5. It holds that $\|\bar{x}'\| \leq \text{OPT}(\text{PRIMAL}(\bar{y}, \mathcal{D})) = \text{OPT}(\text{DUAL}(\bar{y}, \mathcal{D})) \leq (1+\varepsilon)\, h$, where the last inequality is by Lemma 4.39. Since $(\bar{x}', \bar{y}') \in P(\mathcal{D})$, it holds that $\mathcal{E}(\bar{x}') \leq \bar{y}' \in P_{\mathcal{M}}(G)$; thus, $\mathcal{E}(\bar{x}') \in P_{\mathcal{M}}(G)$. For the same reason, we also have $\sum_{C \in \mathcal{C}} \bar{x}'_C \cdot C(i) \geq 1$ for all $i \in I$. Hence, $\bar{x}'$ is a solution for rMLP of value at most $(1+\varepsilon)h$. ◻

Our algorithm for δ-rMLP, given in Algorithm 5, uses Ellipsoid_Q to perform a binary search.

Proof of Lemma 1.8. We show that Algorithm 5 is a polynomial time $(1+3\varepsilon)$-approximation algorithm for rMLP. This immediately implies a PTAS for the MLP problem due to the connection between MLP and rMLP.

Algorithm 5: Matching-LP

Configuration: $\varepsilon, \delta \in (0, 0.1)$
Input : A 2VBP instance (I, v).
Output : A $(1 + O(\varepsilon))$-approximate solution $\bar{x}$ for rMLP.

1 Run a binary search over the range $(\ell, u) = ((1-\delta)|\mathcal{M}^*|, |I|)$: in each iteration call Ellipsoid_Q(h) with $h = \frac{\ell+u}{2}$. If Ellipsoid_Q returned that OPT(rMLP) $> h$ update $\ell = h$; if Ellipsoid_Q returned a solution $\bar{x}$, set $\bar{x}$ to be the best solution and $u = h$. Repeat the process until $u - \ell < \varepsilon$.
2 If $u \neq |I|$, return the best solution found; else, return a vector $\bar{x} \in \{0,1\}^{\mathcal{C}}$ where $\bar{x}_{\{i\}} = 1$ for every $i \in I$ and $\bar{x}_C = 0$ for any other $C \in \mathcal{C}$.

By Lemma 4.40 it holds that OPT(rMLP) $> \ell$ throughout the binary search, and if $u \neq |I|$ then the best solution found $\bar{x}$ satisfies $\|\bar{x}\| \leq (1+\varepsilon)u$ throughout the execution of the binary search. Thus, Algorithm 5 returns a solution $\bar{x}$ satisfying

$$\|\bar{x}\| \leq (1+\varepsilon)u < (1+\varepsilon)(\ell+\varepsilon) < (1+\varepsilon)(\mathrm{OPT(rMLP)}+\varepsilon) \leq (1+3\varepsilon)\mathrm{OPT(rMLP)},$$

where the last inequality holds since OPT(rMLP) ≥ 1 (otherwise $I = \emptyset$ and $\bar{x} = \mathbf{0}$ is an optimal solution). $\square$

It remains to prove Lemma 4.38. Similar to Ellipsoid_Q, the ellipsoid method is applied with an approximate separation oracle. Consider the following family of polytopes. For any $\ell \geq 0$, $\bar{y} \in P_{\mathcal{M}}(G)$ and $\mathcal{D} \subseteq \mathcal{C}$, define

$$\begin{aligned} R(\ell, \bar{y}, \mathcal{D}) &= \left\{ (\bar{\lambda}, \bar{\beta}) \in R(\mathcal{D}) \mid \sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \geq \ell \right\} \\ &= \left\{ (\bar{\lambda}, \bar{\beta}) \in \mathbb{R}^I_{\geq 0} \times \mathbb{R}^E_{\geq 0} \;\middle|\; \begin{array}{ll} & \sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \geq \ell \\ \forall C \in \mathcal{D} \setminus \mathcal{C}_2 : & \sum_{i \in C} \bar{\lambda}_i \leq 1 \\ \forall e \in E,\ C \in S(e) \cap \mathcal{D} : & \sum_{i \in C} \bar{\lambda}_i \leq 1 + \beta_e \end{array} \right\}. \end{aligned} \tag{101}$$

The ellipsoid method is used with polytopes in $R(\ell, \bar{y}, \mathcal{D})$. To derive a separation oracle for $R(\ell, \bar{y}, \mathcal{D})$ we use a PTAS for 2-Dimensional Knapsack (2DK) [FC84]. Using the terminology in this paper, the input for 2DK is a 2VBP instance (S, v), a profit vector $\bar{p} \in \mathbb{R}^S_{\geq 0}$ and a two-dimensional budget $\bar{b} \in \mathbb{R}^2_{\geq 0}$. The objective is to find a subset $W \subseteq S$ of items such that $v(W) = \sum_{i \in W} v(i) \leq \bar{b}$, and $p(W) \equiv \sum_{i \in W} \bar{p}_i$ is maximal. Denote a 2DK instance by $(S, v, \bar{p}, \bar{b})$. We also allow $\bar{p} \in \mathbb{R}^T_{\geq 0}$ where $S \subseteq T$. The separation oracle is given in Algorithm 6. The pseudocode uses $N_G[j] = \{i \in L \mid \{i,j\} \in E\} \cup \{j\}$ to denote the closed neighborhood of $j \in L$ in the δ-matching graph G.

As in the case of Q_separator, we show that R_separator has properties similar to those of a separation oracle.

Lemma 4.41. *On input $(\bar{\lambda}, \bar{\beta}) \in \mathbb{R}^I \times \mathbb{R}^E$, $\bar{y} \in P_{\mathcal{M}}(G)$ and $\ell \geq (1-\delta)|\mathcal{M}^*|$, in polynomial time Algorithm 6 either*

- *returns a separating hyperplane between $R(\ell, \bar{y}, \mathcal{C})$ and $(\bar{\lambda}, \bar{\beta})$, or*

Algorithm 6: R_separator

Input : $(\bar{\lambda}, \bar{\beta}) \in \mathbb{R}^I \times \mathbb{R}^E$, $\bar{y} \in P_{\mathcal{M}}(G)$ and $\ell > (1-\delta)|\mathcal{M}^*|$.

Output: Either a separating hyperplane between $R(\ell, \bar{y}, \mathcal{C})$ and $(\bar{\lambda}, \bar{\beta})$, or $(\bar{\lambda}', \bar{\beta}') \in R\left((1-\frac{\varepsilon}{2})\ell, \bar{y}, \mathcal{C}\right)$.

1 If $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e < \ell$, then return it as the separating hyperplane.

2 Find a $(1-\frac{\varepsilon}{8})$-approximate solution W for the 2DK instance $(I \setminus L, v, \bar{\lambda}, \mathbf{1})$. If $\sum_{i\in W} \bar{\lambda}_i > 1$, return W as a separating hyperplane.

3 **foreach** $j \in L$ **do**

4   Find a $(1-\frac{\varepsilon}{8})$-approximate solution W for the 2DK instance $(I \setminus N_G[j], v, \bar{\lambda}, \mathbf{1} - v(j))$. If $\sum_{i\in W\cup\{j\}} \bar{\lambda}_i > 1$ return $W \cup \{j\}$ as a separating hyperplane.

5 **end**

6 **foreach** $e \in E$ **do**

7   Find a $(1-\frac{\varepsilon}{8})$-approximate solution W for the 2DK instance $(I \setminus L, v, \bar{\lambda}, \mathbf{1} - v(e))$. If $\sum_{i\in W\cup e} \bar{\lambda}_i > 1 + \bar{\beta}_e$ return $W \cup e$ as a separating hyperplane.

8 **end**

9 Notify the ellipsoid method to abort, and return $\left(\left(1-\frac{\varepsilon}{8}\right)\bar{\lambda}, \bar{\beta}'\right)$ where $\bar{\beta}'_e = \min\{2, \bar{\beta}_e\}$ for every $e \in E$.

- *notifies the ellipsoid method to abort and returns* $(\bar{\lambda}', \bar{\beta}') \in R\left(\left(1-\frac{\varepsilon}{2}\right)\ell, \bar{y}, \mathcal{C}\right)$.

Proof. Since 2DK admits a PTAS [FC84], it follows that Algorithm 6 runs in polynomial time.

If $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e < \ell$ then the algorithm returns this inequality as a separating hyperplane in Step 1. This inequality indeed serves as a separating hyperplane by the definition of $R(\ell, \bar{y}, \mathcal{C})$ in (101). Thus, for the remainder of the proof, we may assume that $\sum_{i\in I} \bar{\lambda}_i - \sum_{e\in E} \bar{\beta}_e \cdot \bar{y}_e \geq \ell$.

If the algorithm returns a set W in Step 2, then $W \subseteq I \setminus L$ and $v(W) \leq \mathbf{1}$ as a solution for 2DK. Thus, $W \in \mathcal{C} \setminus \mathcal{C}_2$ and the inequality $\sum_{i\in W} \bar{\lambda}_i > 1$ defines a separating hyperplane by (101) and (99). Hence, for the remainder of the proof we may assume that the algorithm did not return a set in Step 2. This implies that the optimal solution for the 2DK instance $(I \setminus L, v, \bar{\lambda}, \mathbf{1})$ has value at most $\left(1-\frac{\varepsilon}{8}\right)^{-1}$. Since every $C \in \mathcal{C}$ such that $C \subseteq I \setminus L$ is a solution for $(I \setminus L, v, \bar{\lambda}, \mathbf{1})$, it follows that

$$\forall C \in \mathcal{C},\ C \subseteq I \setminus L: \quad \sum_{i\in C} \bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1} \quad . \tag{102}$$

Consider the case in which the algorithm returns the set $W \cup \{j\}$ in Step 4. It holds that $v(W \cup \{j\}) \leq v(W) + v(j) \leq \mathbf{1} - v(j) + v(j) = \mathbf{1}$, as W is a solution for the 2DK instance $(I \setminus N_G[j], v, \bar{\lambda}, \mathbf{1} - v(j))$. Thus, $W \cup \{j\} \in \mathcal{C}$. Suppose, for the sake of contradiction, that $W \cup \{j\} \in \mathcal{C}_2$. Thus, there is some $j' \in W \cap L$ such that $(j, j') \in E$, and we conclude that $W \cap N[j] \neq \emptyset$, contradicting $W \subseteq I \setminus N[j]$ (see Step 3). It therefore holds that $W \cup \{j\} \in \mathcal{C} \setminus \mathcal{C}_2$. Since $\sum_{i\in W\cup\{j\}} \bar{\lambda}_i > 1$, the configuration $W \cup \{j\}$ defines a separating hyperplane, by (101) and (99).

Hence, for the remainder of the proof we may assume that the algorithm did not return a separating hyperplane in Step 4. Let $C \in \mathcal{C} \setminus \mathcal{C}_2$. If $C \subseteq I \setminus L$ then it holds that $\sum_{i\in C} \bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1}$ by (102).

Consider the iteration of the loop in Step 3 in which $j = j^*$, and let W be the set found in this iteration in Step 4. It holds that $C \setminus \{j\}$ is a solution for the 2DK instance $(I \setminus N_G[j], v, \bar{\lambda}, \mathbf{1} - v(j))$; thus, $\sum_{i\in W} \bar{\lambda}_i \geq \left(1-\frac{\varepsilon}{8}\right) \sum_{i\in C\setminus\{j\}} \bar{\lambda}_i$. Since the algorithm did not return $W \cup \{j\}$, we have that

that $\sum_{i\in W\cup\{j\}}\bar{\lambda}_i \leq 1$. Therefore,

$$\sum_{i\in C}\bar{\lambda}_i = \bar{\lambda}_j + \sum_{i\in C\setminus\{j\}}\bar{\lambda}_i \leq \bar{\lambda}_j + \left(1-\frac{\varepsilon}{8}\right)^{-1}\sum_{i\in W}\bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1}\sum_{i\in W\cup\{j\}}\bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1} .$$

Thus,

$$\forall C\in\mathcal{C}\setminus\mathcal{C}_2: \quad \sum_{i\in C}\bar{\lambda}_i \leq \left(1+\frac{\varepsilon}{8}\right)^{-1}. \tag{103}$$

Next, we consider the case in which the algorithms returns the set $W\cup e$ in Step 7. Then $v(W\cup\{e\}) = v(W)+v(e) \leq \mathbf{1}-v(e)+v(e) = \mathbf{1}$ since W is a solution for $(I\setminus L, v, \bar{\lambda}, \mathbf{1}-v(e))$. Hence, $W\cup e\in\mathcal{C}$. It follows that $W\cup\{e\}\in S(e)$. Since $\sum_{i\in W\cup e}\bar{\lambda}_i > 1+\bar{\beta}_e$, it follows that $W\cup e$ defines a separating hyperplane between $(\bar{\lambda},\bar{\beta})$ and $R(\ell,\bar{y},\mathcal{C})$ (by (99) and (101)).

We may therefore assume that the algorithm does not return a set in Step 7 throughout its execution. Let $e^*\in E$ and $C\in S(e^*)$, and consider the iteration of the loop in Step 6 in which $e=e^*$. It holds that $C\setminus e\subseteq I\setminus L$ (otherwise, $v_d(C)>1$ for some $d\in\{1,2\}$) and $v(C\setminus e)\leq \mathbf{1}-v(e)$; thus, $C\setminus e$ is a solution for the 2DK instance $(I\setminus L, v, \bar{\lambda}, \mathbf{1}-v(e))$. Let W be the approximate solution found for $(I\setminus L, v,\bar{\lambda},\mathbf{1}-v(e))$. It then holds that $\sum_{i\in W}\bar{\lambda}_i \geq \left(1-\frac{\varepsilon}{8}\right)\sum_{i\in C\setminus e}\bar{\lambda}_e$. Also, since we assume that the algorithm does not return a set in Step 7, it holds that $\sum_{i\in W\cup e}\bar{\lambda}\leq 1+\beta_e$. Therefore, we have that

$$\sum_{i\in C}\bar{\lambda}_i = \sum_{i\in e}\bar{\lambda}_i + \sum_{i\in C\setminus e}\bar{\lambda}_i \leq \sum_{i\in e}\bar{\lambda}_i + \left(1-\frac{\varepsilon}{8}\right)^{-1}\sum_{i\in W}\bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1}\sum_{i\in W\cup e}\bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1}(1+\beta_e) . \tag{104}$$

Let $e=\{j_1,j_2\}$. Then $\{j_1\},\{j_2\},C\setminus e\in\mathcal{C}\setminus\mathcal{C}_2$. Therefore, by (103),

$$\sum_{i\in C}\bar{\lambda}_i \leq \bar{\lambda}_{j_1}+\bar{\lambda}_{j_2}+\sum_{i\in C\setminus e}\bar{\lambda}_i \leq 3\left(1+\frac{\varepsilon}{8}\right)^{-1} . \tag{105}$$

By (104) and (105), we have

$$\forall e\in E,\ C\in S(e): \quad \sum_{i\in C}\bar{\lambda}_i \leq \left(1-\frac{\varepsilon}{8}\right)^{-1}(1+\min\{\bar{\beta}_e,2\}) = \left(1-\frac{\varepsilon}{8}\right)^{-1}(1+\bar{\beta}'_e) . \tag{106}$$

By (103) and (106) it holds that $\left(\left(1-\frac{\varepsilon}{8}\right)\bar{\lambda},\bar{\beta}'\right)\in R(\mathcal{C})$. Furthermore,

$$\begin{aligned}\sum_{i\in I}\left(1-\frac{\varepsilon}{8}\right)\bar{\lambda}_i - \sum_{e\in E}\bar{\beta}'_e\cdot\bar{y}_e &= \left(1-\frac{\varepsilon}{8}\right)\cdot\left(\sum_{i\in I}\bar{\lambda}_i - \sum_{e\in E}\bar{\beta}'_e\cdot\bar{y}_e\right) - \frac{\varepsilon}{8}\sum_{e\in E}\bar{\beta}'_e\cdot\bar{y}_e\\ &\geq \left(1-\frac{\varepsilon}{8}\right)\cdot\left(\sum_{i\in I}\bar{\lambda}_i - \sum_{e\in E}\bar{\beta}_e\cdot\bar{y}_e\right) - \frac{\varepsilon}{4}\sum_{e\in E}\bar{y}_e\\ &\geq \left(1-\frac{\varepsilon}{8}\right)\ell - \frac{\varepsilon}{4}\frac{\ell}{1-\delta}\\ &\geq \left(1-\frac{\varepsilon}{2}\right)\ell .\end{aligned}$$

The first inequality holds since $\bar{\beta}'_e=\min\{\bar{\beta}_e,2\}$, the second inequality uses $\sum_{e\in E}\bar{y}_e\leq|\mathcal{M}^*|<\frac{\ell}{1-\delta}$ due to (100). Thus, $\left(\left(1-\frac{\varepsilon}{8}\right)\bar{\lambda},\bar{\beta}'\right)\in R\left(\left(1-\frac{\varepsilon}{2}\right)\ell,\bar{y},\mathcal{C}\right)$. ☐

The facet complexity of $R(\ell, \bar{y}, \mathcal{D})$ can be trivially bounded by (99), as stated in the next lemma (we omit the proof).

Lemma 4.42. *There is a polynomial p_2 (independent of the instance (I, v)) such that for any $\mathcal{D} \subseteq \mathcal{C}$, $\bar{y} \in P_{\mathcal{M}}$ and $\ell \geq 0$ the facet complexity of $R(\ell, \bar{y}, \mathcal{D})$ is at most $p_2(|I| + \langle \bar{y} \rangle + \langle \ell \rangle)$.*

Algorithm 7 uses the ellipsoid method with $\mathsf{R\_separator}$ as the separation oracle.

Algorithm 7: $\mathsf{Ellipsoid\_R}$

Input : $\bar{y} \in P_{\mathcal{M}}(G)$ and $h > (1-\delta)|\mathcal{M}^*|$

Output: Either a subset $\mathcal{D} \subseteq \mathcal{C}$ such that $\mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{D})) \leq (1+\varepsilon)h$ or a point $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$ such that $\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e > h$.

Run $\mathsf{Ellipsoid}$ (Proposition 4.33) with $n = |I| + |E|$, $\varphi = p_2(|I| + \langle \bar{y} \rangle + \langle \ell \rangle)$ and $\mathsf{R\_separator}$ as the separation oracle, where $\mathsf{R\_separator}$ is used with $\bar{y}$ and $\ell = \frac{h}{1 - \frac{3\varepsilon}{4}}$.

if *the ellipsoid method returned the polytope is empty* **then**

  Let $\mathcal{D}$ be the set of configurations returned by $\mathsf{R\_separator}$ as a separating hyperplanes throughout the execution of the ellipsoid method. Return $\mathcal{D}$.

else

  `// This only happens if R_separator aborted the ellipsoid method.`

  Return $(\bar{\lambda}, \bar{\beta})$, where $(\bar{\lambda}, \bar{\beta}) \in \left(\left(1 - \frac{\varepsilon}{2}\right) \ell, \bar{y}, \mathcal{C}\right)$ is the value returned by $\mathsf{R\_separator}$.

end

Proof of Lemma 4.38. Note that $\mathsf{Ellipsoid\_R}$ runs in polynomial time. Furthermore, $\ell > h > (1-\delta)|\mathcal{M}^*|$. Thus, $\mathsf{R\_separator}$ is used with parameters that match the conditions of Lemma 4.41.

Consider the execution of Algorithm 7. If the ellipsoid method returns that the polytope is empty then all separating hyperplanes returned by $\mathsf{Ellipsoid\_R}$ are also separating hyperplanes with respect to the polytope $R(\ell, \bar{y}, \mathcal{D})$. Thus, it must hold that $R(\ell, \bar{y}, \mathcal{D}) = \emptyset$. This implies that $\mathrm{OPT}(\mathrm{DUAL}(\bar{y}, \mathcal{D})) \leq \ell = \frac{h}{1 - \frac{3\varepsilon}{4}} \leq (1+\varepsilon)h$. Since the execution of the ellipsoid is of polynomial time, it follows that $|\mathcal{D}|$ is also polynomial.

If the ellipsoid method was aborted, then by Lemma 4.41 it holds that $(\bar{\lambda}, \bar{\beta}) \in \left(\left(1 - \frac{\varepsilon}{2}\right) \ell, \bar{y}, \mathcal{C}\right)$. By (101) we have that $(\bar{\lambda}, \bar{\beta}) \in R(\mathcal{C})$, and

$$\sum_{i \in I} \bar{\lambda} - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \geq \left(1 - \frac{\varepsilon}{2}\right) \ell = \left(1 - \frac{\varepsilon}{2}\right) \frac{h}{1 - \frac{3\varepsilon}{4}} > h \ . \qquad \square$$

5 Basic Probabilistic Tools

In this section we prove Lemmas 2.4 and 2.6; the probabilistic lemmas which are used both in Section 3 and Section 4. The proof of Lemma 2.4 follows from an iterative application of Lemma 2.3. Lemma 2.6 is an application of Lemma 2.4.

We begin with the following technical lemma.

Lemma 5.1. *Let $j \in \{0, 1, \ldots, k-1\}$ and $t > 0$. Also, let $\bar{u} \in \mathbb{R}^I_{\geq 0}$ be an $\mathcal{F}_j$-measurable random vector. Then,*

$$\Pr\left(\bar{u} \cdot \mathbb{1}_{S_{j+1}} - (1-\delta)\bar{u} \cdot \mathbb{1}_{S_j} > t \cdot \mathsf{tol}(\bar{u})\right) \ \leq \ \exp\left(-\frac{2 \cdot t^2}{\mathrm{OPT}}\right) \ .$$

Proof. Let A be the set of possible values the random vector $\bar{u}$ can take, that is, $A = \{\bar{u}(\omega) \mid \omega \in \Omega\}$. Since Ω is finite, it holds that A is also finite.

For any $S \subseteq I$, $\rho \in \{1, \ldots, \text{OPT}\}$ and $\bar{a} \in A$ define $f_{S,\rho,\bar{a}} : \mathcal{C}^{\text{OPT}} \to \mathbb{R}$ by

$$f_{S,\rho,\bar{a}}(C_1, \ldots, C_{\text{OPT}}) = \begin{cases} \frac{1}{\mathsf{tol}(\bar{a})} \cdot \bar{a} \cdot \mathbb{1}_{S \setminus (\bigcup_{\ell=1}^{\rho} C_\ell)}, & \mathsf{tol}(\bar{a}) \neq 0, \\ 0, & \text{otherwise} \end{cases} .$$

Also, define $D = \{f_{S,\rho,\bar{a}} \mid S \subseteq I, \rho \in \{1, \ldots, \text{OPT}\}, \bar{a} \in A\}$. It can be easily verified that D is finite.

Let $f_{S,\rho,\bar{a}} \in D$, $(C_1, \ldots, C_{\text{OPT}})$, $(C'_1, \ldots, C'_{\text{OPT}}) \in \mathcal{C}^{\text{OPT}}$ and $r \in [\text{OPT}]$ such that $C_\ell = C'_\ell$ for $\ell = 1, \ldots, r-1, r+1, \ldots, \text{OPT}$. If $\mathsf{tol}(\bar{a}) = 0$ or $r > \rho$, then

$$\left| f_{S,\rho,\bar{a}}(C_1, \ldots, C_{\text{OPT}}) - f_{S,\rho,\bar{a}}(C'_1, \ldots, C'_{\text{OPT}}) \right| = 0 \ .$$

Otherwise, let $T = \bigcup_{\ell \in [\rho] \setminus \{r\}} C_\ell = \bigcup_{\ell \in [\rho] \setminus \{r\}} C'_\ell$. Then

$$\begin{aligned}
\left| f_{S,\rho,\bar{a}}(C_1, \ldots, C_{\text{OPT}}) - f_{S,\rho,\bar{a}}(C'_1, \ldots, C'_{\text{OPT}}) \right| &= \left| \frac{1}{\mathsf{tol}(\bar{a})} \cdot \bar{a} \left(\mathbb{1}_{S \setminus T \setminus C_r} - \mathbb{1}_{S \setminus T \setminus C'_r} \right) \right| \\
&= \left| \frac{1}{\mathsf{tol}(\bar{a})} \left(\sum_{i \in (S \cap C'_r) \setminus (C_r \cup T)} \bar{a}_i - \sum_{i \in (S \cap C_r) \setminus (C'_r \cup T)} \bar{a}_i \right) \right| \\
&\leq \frac{1}{\mathsf{tol}(\bar{a})} \cdot \max \left\{ \sum_{i \in (S \cap C'_r) \setminus (C_r \cup T)} \bar{a}_i, \sum_{i \in (S \cap C_r) \setminus (C'_r \cup T)} \bar{a}_i \right\} \\
&\leq \frac{1}{\mathsf{tol}(\bar{a})} \cdot \mathsf{tol}(\bar{a}) \leq 1 \ .
\end{aligned}$$

The second equality holds, as $S \setminus T \setminus C_r \setminus (S \setminus T \setminus C'_r) = (S \cap C'_r) \setminus (C_r \cup T)$ and symmetrically $S \setminus T \setminus C'_r \setminus (S \setminus T \setminus C_r) = (S \cap C_r) \setminus (C'_r \cup T)$. Thus, $f_{S,\rho,\bar{a}}$ is of 1-bounded difference.

Define a random function $g = f_{S_j, \rho_{j+1}, \bar{u}}$. Since S_j, ρ_{j+1} and $\bar{u}$ are $\mathcal{F}_j$-measurable, it follows that g is $\mathcal{F}_j$-measurable. By definition of g, we have

$$\mathsf{tol}(\bar{u}) \cdot g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) = \bar{u} \cdot \mathbb{1}_{S_j \setminus \bigcup_{\ell=1}^{\rho_{j+1}} C_\ell^{j+1}} = \bar{u} \cdot \mathbb{1}_{S_{j+1}} \ .$$

Furthermore,

$$\begin{aligned}
\mathbb{E}[\mathsf{tol}(\bar{a}) \cdot g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) \mid \mathcal{F}_j] &= \mathbb{E}[\bar{u} \cdot \mathbb{1}_{S_{j+1}} \mid \mathcal{F}_j] = \sum_{i \in I} \bar{u}_i \cdot \Pr(i \in S_{j+1} \mid \mathcal{F}_j) \\
&\leq (1-\delta) \sum_{i \in I} \bar{u}_i \cdot \mathbb{1}_{i \in S_j} = (1-\delta) \cdot \bar{a} \cdot \mathbb{1}_{S_j},
\end{aligned}$$

where the inequality holds by Lemma 2.1. Therefore,

$$\begin{aligned}
&\Pr\left(\bar{u} \cdot \mathbb{1}_{S_j+1} - (1-\delta)\bar{u} \cdot \mathbb{1}_{S_j} > t \cdot \mathsf{tol}(\bar{u}) \right) \\
&\qquad \leq \Pr\left(g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) - \mathbb{E}[g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1} \mid \mathcal{F}_j)] > t \right) \leq \exp\left(-\frac{2 \cdot t^2}{\text{OPT}} \right),
\end{aligned}$$

where the last inequality is by Lemma 2.3. $\square$

We use Lemma 5.1 as part of the proof of Lemma 2.4

Proof of Lemma 2.4. We note that

$$
\begin{aligned}
&\Pr\left(\exists r\in\{j,\ldots,k\}:\ \bar{u}\cdot\mathbb{1}_{S_r}-(1-\delta)^{r-j}\cdot\bar{u}\cdot\mathbb{1}_{S_j}>t\cdot\mathsf{tol}(\bar{u})\right)\\
=\ &\Pr\left(\exists r\in\{j,\ldots,k\}:\ \sum_{\ell=j+1}^{r}\left(\bar{u}\cdot\mathbb{1}_{S_\ell}-(1-\delta)\cdot\bar{u}\cdot\mathbb{1}_{S_{\ell-1}}\right)\cdot(1-\delta)^{r-\ell}>t\cdot\mathsf{tol}(\bar{u})\right)\\
\leq\ &\Pr\left(\exists r\in\{j+1,\ldots,k\},\ \ell\in\{j+1,\ldots,r\}:\ \left(\bar{u}\cdot\mathbb{1}_{S_\ell}-(1-\delta)\cdot\bar{u}\cdot\mathbb{1}_{S_{\ell-1}}\right)\cdot(1-\delta)^{r-\ell}>\frac{t}{r-j}\cdot\mathsf{tol}(\bar{u})\right)\\
\leq\ &\Pr\left(\exists\ \ell\in\{j+1,\ldots,k\}:\ \bar{u}\cdot\mathbb{1}_{S_\ell}-(1-\delta)\cdot\bar{u}\cdot\mathbb{1}_{S_{\ell-1}}>\frac{t}{k}\cdot\mathsf{tol}(\bar{u})\right)\\
\leq\ &\sum_{\ell=j+1}^{k}\Pr\left(\bar{u}\cdot\mathbb{1}_{S_\ell}-(1-\delta)\cdot\bar{u}\cdot\mathbb{1}_{S_{\ell-1}}>\frac{t}{k}\cdot\mathsf{tol}(\bar{u})\right)\\
\leq\ &k\cdot\exp\left(-\frac{2\cdot\left(\frac{t}{k}\right)^2}{\mathrm{OPT}}\right)\\
\leq\ &\delta^{-2}\exp\left(-\frac{2\cdot\delta^4\cdot t^2}{\mathrm{OPT}}\right)\quad.
\end{aligned}
$$

The first inequality holds, since if a sum of n variables is greater than T there most be a variable with value greater than $\frac{T}{n}$. The fourth inequality is by Lemma 5.1, and the last inequality uses $k\leq\delta^{-2}$. □

Lemma 2.6 is a simple application of Lemma 2.4.

Proof of Lemma 2.6. Define $\bar{u}\in[0,1]^I$ by $\bar{u}_i=\sum_{t=1}^{d}v_t(i)$. For any $C\in\mathcal{C}$ it holds that $\sum_{i\in C}\bar{u}_i=\sum_{t=1}^{d}v_t(C)\leq d$, therefore $\mathsf{tol}(\bar{u})\leq d$. Furthermore, there is partition $(Q_1,\ldots,Q_{\mathrm{OPT}})$ of I such that Q_ℓ is a configuration for $\ell=1,\ldots,\mathrm{OPT}$. Therefore,

$$\bar{u}\cdot\mathbb{1}_{S_0}\leq\bar{u}\cdot\mathbb{1}_I=\sum_{\ell=1}^{\mathrm{OPT}}\bar{u}\cdot\mathbb{1}_{Q_\ell}\leq\mathrm{OPT}\cdot\mathsf{tol}(\bar{u})\leq d\cdot\mathrm{OPT}\quad. \tag{107}$$

Recall that ρ^* is the number of configurations used by First-Fit in Line 6 of Algorithm 1. Using Lemma 2.5, we have

$$
\begin{aligned}
\Pr(\rho^*>8\cdot d\cdot\delta\cdot\mathrm{OPT}+1)&\leq\Pr\left(\sum_{t=1}^{d}v_t(S_k)>4\cdot d\cdot\delta\cdot\mathrm{OPT}\right)\\
&\leq\Pr(\bar{u}\cdot\mathbb{1}_{S_k}>4\cdot d\cdot\delta\cdot\mathrm{OPT})\\
&\leq\Pr\left(\bar{u}\cdot\mathbb{1}_{S_k}-(1-\delta)^k\cdot\bar{u}\cdot\mathbb{1}_{S_0}>3\cdot d\cdot\delta\cdot\mathrm{OPT}\right)\\
&\leq\Pr\left(\exists r\in\{0,\ldots,k\}:\ \ \bar{u}\cdot\mathbb{1}_{S_r}-(1-\delta)^r\cdot\bar{u}\cdot\mathbb{1}_{S_0}>\mathsf{tol}(\bar{u})\cdot\delta\cdot\mathrm{OPT}\right)\\
&\leq\delta^{-2}\cdot\exp\left(-\frac{2\cdot\delta^4\cdot\delta^2\cdot\mathrm{OPT}^2}{\mathrm{OPT}}\right)\\
&\leq\delta^{-2}\cdot\exp\left(-\delta^7\cdot\mathrm{OPT}\right)\quad.
\end{aligned}
$$

The third inequality uses (107) and $(1-\delta)^k\leq\delta$. The fifth inequality is by Lemma 2.4. Hence, $\Pr(\rho^*\leq8\cdot d\cdot\delta\cdot\mathrm{OPT}+1)\geq1-\delta^{-2}\cdot\exp(-\delta^7\cdot\mathrm{OPT})$. □

6 Discussion

In this paper we showed that a simple iterative randomized rounding scheme (Algorithm 1) improves the state-of-the-art algorithms for d-DIMENTIONAL VECTOR BIN PACKING, for any $d > 3$. We also showed that Algorithm 1 outperforms any algorithm within the *Round&Approx* framework of Bansal et al. [BCS10]. Slight modifications in this algorithm to include an initial matching phase (Algorithm 2) led to an algorithm that yields an asymptotic $\left(\frac{4}{3}+\varepsilon\right)$-approximation for 2-DIMENTIONAL VECTOR BIN PACKING, improving upon the $\left(\frac{3}{2}+\varepsilon\right)$-approximation algorithm of Bansal et al. [BEK16]. To the best of our knowledge, we use here for the first time iterative *randomized* rounding in the context of BIN PACKING problems.

For arbitrary $d > 2$ we applied a fairly simple analysis of Algorithm 1, which leaves much room for improvement. Our analysis of Algorithm 2 is the result of multiple back-and-forth steps which led to new insights on the stochastic process generated by randomized rounding, and on structural properties of dVBP which proved useful in the analysis. The matching subroutine in Algorithm 2 was introduced as part of this process. While this led to a significantly better asymptotic approximation ratio for $d = 2$, our analysis for this case is more complex.

We note that many of the ideas used in the analysis for $d = 2$ can be easily incorporated into the analysis for $d > 2$. For example, the sets T_j (defined in (12)) used in the proof of Theorem 1.5 are analogous to touched configurations in the analysis of Section 4.1.1. While the analysis for $d = 2$ considers the set T_j for every iteration j and attempts to exploit it to improve the approximation ratio, the analysis for arbitrary d only considers the set T_{j_1} for a specific value of j_1.

As part of the analysis of Algorithm 2 we introduced a structural property for 2VBP (Lemma 4.2) which combines ideas of Bansal et al. [BEK16] and Fairstein et al. [FKS21]. Intuitively, it should be possible to extend the lemma to arbitrary $d > 2$. While the rounding scheme presented in the proof of Lemma 4.2 can be extended to $d > 2$, the Small Items Refinement (Lemma 4.23) is tailored to the two-dimensional case.

The basic idea behind Algorithm 1 is that covering items with some fixed probability via iterative randomized rounding requires sampling fewer configurations, in comparison to non-iterative rounding. In our proofs we used structural properties of dVBP (e.g, Lemmas 3.5 and 4.2) to formalize this basic idea. Intuitively, the same basic idea should also work for other BIN PACKING problems, such as GEOMETRIC 2-DIMENSIONAL BIN PACKING [BK14] and GENERALIZED MULTIDIMENSIONAL BIN PACKING [KSS21], for which the state-of-the-art algorithms use the *Round&Approx* framework. Formalizing this intuition requires an analog of the structural properties for each of these BIN PACKING variants. We note that, even without a tailored structural property, following the outline of the proof of Theorem 3.2, it can be easily shown that a simple adaptation of Algorithm 1 yields an asymptotic approximation ratio which is at least as good as the ratio of any *Round&Approx* algorithm for GEOMETRIC 2-DIMENSIONAL BIN PACKING [BK14] and for GENERALIZED MULTIDIMENSIONAL BIN PACKING [KSS21].

Algorithm 1 can be used also to simplify existing results. For example, in Lemma 1.6 we showed the algorithm is an AFPTAS for BIN PACKING. We conjecture that the algorithm is also an AFPTAS for BIN PACKING WITH CARDINALITY CONSTRAINTS [EL10].

Finally, the number of configurations sampled in each iteration of Line 1 in Algorithm 1 was selected arbitrarily for an easier analysis. One may consider selecting a *single* configuration per iteration. We believe that such modification is unlikely to yield a better approximation ratio, but rather make the analysis more complicated. A main cause for complication here is that the vanilla form of McDiarmid's concentration bound [McD89] cannot be used, due to stronger dependencies between the sampled configurations.

References

[ADGH18] Roberto Aringhieri, Davide Duma, Andrea Grosso, and Pierre Hosteins. Simple but effective heuristics for the 2-constraint bin packing problem. *J. Heuristics*, 24(3):345–357, 2018.

[Ban14] Nikhil Bansal. New developments in iterated rounding (invited talk). In *Proc. FSTTCS 2014*, volume 29 of *Leibniz Int. Proc. Informatics*, pages 1–10, 2014.

[Ban19] Nikhil Bansal. On a generalization of iterated and randomized rounding. In *Proc. STOC 2019*, pages 1125–1135, 2019.

[BCS10] Nikhil Bansal, Alberto Caprara, and Maxim Sviridenko. A new approximation method for set covering problems, with applications to multidimensional bin packing. *SIAM J. Comput.*, 39(4):1256–1278, 2010.

[BEK16] Nikhil Bansal, Marek Eliáš, and Arindam Khan. Improved approximation for vector bin packing. In *Proc. SODA 2016*, pages 1561–1579, 2016.

[BEK21] Nikhil Bansal, Marek Eliáš, and Arindam Khan. Personal Communication, 2021.

[BGRS13] Jarosław Byrka, Fabrizio Grandoni, Thomas Rothvoß, and Laura Sanità. Steiner tree approximation via iterative randomized rounding. *J. ACM*, 60(1):1–33, 2013.

[BK14] Nikhil Bansal and Arindam Khan. Improved approximation algorithm for two-dimensional bin packing. In *Proc. SODA 2014*, pages 13–25, 2014.

[CCG+13] Edward G Coffman, János Csirik, Gábor Galambos, Silvano Martello, and Daniele Vigo. Bin packing approximation algorithms: survey and classification. In *Handbook of combinatorial optimization*, pages 455–531. 2013.

[CHP05] Soo Y Chang, Hark-Chin Hwang, and Sanghyuck Park. A two-dimensional vector packing model for the efficient use of coil cassettes. *Comput. Oper. Res.*, 32(8):2051–2058, 2005.

[CK04] Chandra Chekuri and Sanjeev Khanna. On multidimensional packing problems. *SIAM J. Comput.*, 33(4):837–851, 2004.

[CKPT17] Henrik I Christensen, Arindam Khan, Sebastian Pokutta, and Prasad Tetali. Approximation and online algorithms for multidimensional bin packing: A survey. *Comput. Sci. Rev.*, 24:63–79, 2017.

[CLRS01] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms.* MIT Press, third edition, 2001.

[CT97] Yuan Shih Chow and Henry Teicher. *Probability theory: independence, interchangeability, martingales.* Springer Science & Business Media, 1997.

[CVZ11] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Multi-budgeted matchings and matroid intersection via dependent rounding. In *Proc. SODA 2011*, pages 1080–1097, 2011.

[DIM16] Maxence Delorme, Manuel Iori, and Silvano Martello. Bin packing and cutting stock problems: Mathematical models and exact algorithms. *Europ. J. Oper. Res.*, 255(1):1–20, 2016.

[EL10] Leah Epstein and Asaf Levin. AFPTAS results for common variants of bin packing: A new method for handling the small items. *SIAM J. Optim*, 20(6):3121–3145, 2010.

[FC84] Alan M Frieze and Michael RB Clarke. Approximation algorithms for the m-dimensional 0-1 knapsack problem: worst-case and probabilistic analyses. *Europ. J. Oper. Res.*, 15(1):100–109, 1984.

[FKS21] Yaron Fairstein, Ariel Kulik, and Hadas Shachnai. Modular and submodular optimization with multiple knapsack constraints via fractional grouping. In *Proc. ESA 2021*, Leibniz Int. Proc. Informatics, pages 41:1–41:16, 2021.

[FL81] W Fernandez de la Vega and George S Lueker. Bin packing can be solved within $1 + \varepsilon$ in linear time. *Combinatorica*, 1(4):349–355, 1981.

[GGJY76] Michael R Garey, Ronald L Graham, David S Johnson, and Andrew Chi-Chih Yao. Resource constrained scheduling as generalized bin packing. *J. Comb. Theory, Ser. A*, 21(3):257–298, 1976.

[GLS81] Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[GLS88] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer-Verlag, 1988.

[HR17] Rebecca Hoberg and Thomas Rothvoss. A logarithmic additive integrality gap for bin packing. In *Proc. SODA 2017*, pages 2616–2625, 2017.

[HS22] Mhand Hifi and Shohre Sadeghsa. An iterative randomized rounding algorithm for the k-clustering minimum completion problem with an application in telecommunication field. In *Intelligent Computing: Proceedings of the 2021 Computing Conference, Volume 1*, pages 410–422, 2022.

[IM20] Sungjin Im and Benjamin Moseley. Fair scheduling via iterative quasi-uniform sampling. *SIAM J. Comput.*, 49(3):658–680, 2020.

[Joh16] David S. Johnson. Vector bin packing. In *Encyclopedia of Algorithms*, pages 2319–2323, 2016.

[KK82] Narendra Karmarkar and Richard M. Karp. An efficient approximation scheme for the one-dimensional bin-packing problem. In *Proc. FOCS 1982*, pages 312–320, 1982.

[KK03] Hans Kellerer and Vladimir Kotov. An approximation algorithm with absolute worst-case performance ratio 2 for two-dimensional vector packing. *Oper. Res. Lett.*, 31(1):35–41, 2003.

[KSS21] Arindam Khan, Eklavya Sharma, and KVN Sreenivas. Geometry meets vectors: Approximation algorithms for multidimensional packing. Technical report, 2021. https://arxiv.org/abs/2106.13951.

[LRS11] Lap Chi Lau, Ramamoorthi Ravi, and Mohit Singh. *Iterative methods in combinatorial optimization*, volume 46. Cambridge University Press, 2011.

[McD89] Colin McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

[MT06] Michele Monaci and Paolo Toth. A set-covering-based heuristic approach for bin-packing problems. *INFORMS J. Comput.*, 18(1):71–85, 2006.

[PTUW11] Rina Panigrahy, Kunal Talwar, Lincoln Uyeda, and Udi Wieder. Heuristics for vector bin packing. Technical report, 2011. https://www.microsoft.com/en-us/research/wp-content/uploads/2011/01/VBPackingESA11.p

[Ray21] Arka Ray. There is no APTAS for 2-dimensional vector bin packing: Revisited. Technical report, 2021. https://arxiv.org/abs/2104.13362.

[Rot17] Thomas Rothvoß. The matching polytope has exponential extension complexity. *J. ACM*, 64(6):1–19, 2017.

[San22] Sai Sandeep. Almost optimal inapproximability of multidimensional packing problems. In *Proc. FOCS 2021*, pages 245–256, 2022.

[Sch03] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.

[Spi94] Frits CR Spieksma. A branch-and-bound algorithm for the two-dimensional vector packing problem. *Comput. Oper. Res.*, 21(1):19–25, 1994.

[TS19] Asser N Tantawi and Malgorzata Steinder. Autonomic cloud placement of mixed workload: An adaptive bin packing algorithm. In *Proc. ICAC 2019*, pages 187–193, 2019.

[Vaz01] Vijay V Vazirani. *Approximation algorithms*. Springer-Verlag, 2001.

[WLLH20] Lijun Wei, Minghui Lai, Andrew Lim, and Qian Hu. A branch-and-price algorithm for the two-dimensional vector packing problem. *Europ. J. Oper. Res.*, 281(1):25–35, 2020.

[Woe97] Gerhard J Woeginger. There is no asymptotic PTAS for two-dimensional vector packing. *Inf. Proc. Lett.*, 64(6):293–297, 1997.

[WS11] David P Williamson and David B Shmoys. *The design of approximation algorithms*. Cambridge University Press, 2011.

[YG12] Yonghong Yu and Yang Gao. Constraint programming-based virtual machines placement algorithm in datacenter. In *Proc. IIP 2012*, pages 295–304, 2012.

A The Flaw in Bansal, Eliáš and Khan [BEK16]

The flaw we found in the work of Bansal et al. [BEK16] is in the proof of Theorem 6.1. The theorem refers to properties of the residual items after sampling configurations using a solution for the Configuration-LP. The proof of the theorem relies on McDiarmid's bound, given as Lemma 6.1 in [BEK16]. The flaw is in the use of Lemma 6.1, affecting the correctness of the analysis of the asymptotic approximation guarantees of Algorithm 3 and Algorithm 4 in [BEK16]. We refer below to the third paragraph in the left column of page 1575 in [BEK16] (starting with "We now consider the small items"). As some of the ingredients in the proof of Theorem 6.1 are missing, we expand steps and add details where necessary, while keeping the deviation from [BEK16] to a minimum.

Using the notation of [BEK16], let $\rho > 1$, let $\bar{x}$ be a solution for the Configuration-LP (1) of the dVBP instance (I, v), and let $X_1, ..., X_r \sim \bar{x}$ be a tuple of $t = \lceil \rho \cdot z^* \rceil$ random configurations distributed by $\bar{x}$, where $z^* = \|\bar{x}\|$. Also, define $J = I \setminus (\bigcup_{\ell=1}^{r} X_\ell)$ to be the items *not* selected by the sampled configurations $X_1, \ldots, X_r$.

For $j = (h_1, \ldots, h_d) \in [0,1]^d$, $\mathcal{S}_j \subseteq I$ is a set of items such that $v_k(i) \leq h_k$ for all $i \in \mathcal{S}_j$ and $k = 1, \ldots, d$. The set $\mathcal{S}_j$ represents a class of small items. Bansal et al. [BEK16] define functions $f^k_{\mathcal{S}_j} : \mathcal{C}^r \to \mathbb{R}$ by

$$f^k_{\mathcal{S}_j}(C_1, \ldots, C_r) = \sum_{i \in \mathcal{S}_j \setminus (\bigcup_{\ell=1}^{r} C_\ell)} v_k(i) \cdot \frac{1}{h_k} \tag{108}$$

for $k = 1, \ldots, d$. The definition in [BEK16] is: "Let function $f^k_{\mathcal{S}_j}$ be $\sum_{i \in \mathcal{S}_j \cap J} v_k(i) \cdot \frac{1}{h_k}$" (up to a minor adaptation to our slightly different notation), which we can only interpret as (108) due to the subsequent use of $f^k_{\mathcal{S}_j}$ in [BEK16] as a function whose domain is a tuple of configurations, and since

$$f^k_{\mathcal{S}_j}(X_1, \ldots, X_r) = \frac{1}{h_k} \cdot \sum_{i \in \mathcal{S}_j \setminus (\bigcup_{\ell=1}^{r} X_\ell)} v_k(i) = \frac{1}{h_k} \sum_{i \in \mathcal{S}_j \cap J} v_k(i).$$

To use Lemma 6.1 the authors of [BEK16] attempt to show that $f^k_{\mathcal{S}_j}$ is of 1-bounded difference (see the definition in Section 2 of the preset paper) for $k = 1, \ldots, d$. To this end, they consider $\ell^* \in \{1, \ldots, r\}$ and two vectors $x = (C_1, \ldots, C_r) \in \mathcal{C}^r$ and $x' = (C'_1, \ldots, C'_r) \in \mathcal{C}^r$ such that $C_\ell = C'_\ell$ for $\ell \in \{1, \ldots, r\} \setminus \ell^*$. That is, x and x' differ only in one coordinate. Subsequently, the authors state the following:

$$\begin{aligned} & f^k_{\mathcal{S}_j}(C_1, \ldots, C_k) - f^k_{\mathcal{S}_j}(C'_1, \ldots, C'_k) \\ \leq \; & \max \left\{ \sum_{i \in \mathcal{S}_j \cap C_{\ell*}} v_k(i) \cdot \frac{1}{h_k}, \sum_{i \in \mathcal{S}_j \cap C'_{\ell*}} v_k(i) \cdot \frac{1}{h_k} \right\} \\ \leq \; & \frac{1}{h_k} \cdot h_k \; \leq \; 1. \end{aligned} \tag{109}$$

The second inequality (marked is red) is incorrect. With no explanation for this inequality, it appears that Bansal et al. [BEK16] assumed that $v_k(\mathcal{S}_j \cap C) \leq h_k$ for any $C \in \mathcal{C}$. However, there may be $C \in \mathcal{C}$ such that $v_k(\mathcal{S}_j \cap C) = 1$. For example, suppose that $h_k = \frac{1}{10}$, and let $v_k(i) = h_k$ and $v_{k'}(i) = 0$ for every $i \in \mathcal{S}_j$ and $k' \in \{1, \ldots, d\} \setminus \{k\}$. Then a configuration C containing 10 items from $\mathcal{S}_j$ satisfies $v_k(\mathcal{S}_j \cap C) = 1 > h_k$.

In the setting of the proof of Theorem 6.1 of [BEK16], the items in $\mathcal{S}_j$ are assigned to configurations $C^*_1, \ldots, C^*_m$ in a specific solution. Indeed, it holds that $v_k(C^*_\ell \cap \mathcal{S}_j) \leq h_k$ for $\ell = 1, \ldots, m$, and we believe this led the authors of [BEK16] to the conclusion that $v_k(C \cap \mathcal{S}_j) \leq h_k$ for every configuration $C \in \mathcal{C}$, and hence to the flawed inequality in (109).

Thus, the proof that $f^k_{\mathcal{S}_j}$ is of 1-bounded difference is incorrect, and the subsequent use of Lemma 6.1 fails.

A correct version of (109) is

$$\begin{aligned} & f^k_{\mathcal{S}_j}(C_1, \ldots, C_k) - f^k_{\mathcal{S}_j}(C'_1, \ldots, C'_k) \\ \leq \quad & \max\left\{ \sum_{i \in \mathcal{S}_j \cap C_{\ell^*}} v_k(i) \cdot \frac{1}{h_k}, \sum_{i \in \mathcal{S}_j \cap C'_{\ell^*}} v_k(i) \cdot \frac{1}{h_k} \right\} \\ \leq \; & \frac{1}{h_k}. \end{aligned} \tag{110}$$

However, this inequality only shows that $f^k_{\mathcal{S}_j}$ is of $\frac{1}{h_k}$-bounded difference. As $\frac{1}{h_k}$ may be large (for example, it may be that $\frac{1}{h_k} = (\mathrm{OPT}(I))^3$), the concentration bound which can be derived from (110) is too weak to complete the proof.

Theorem 6.1 of [BEK16] is a central component in the proofs of the asymptotic $\left(1 + \ln\left(\frac{3}{2}\right) + \varepsilon\right)$-approximation for 2VBP and of the asymptotic $\left(1.5 + \ln\left(\frac{d+1}{2}\right) + \varepsilon\right)$-approximation for dVBP. By the above, the two results are incorrect.