An Asymptotic \((\frac{4}{3} + \varepsilon)\)-Approximation for the 2-Dimensional Vector Bin Packing Problem

Ariel Kulik\(^*\)  
Matthias Mnich\(^†\)  
Hadas Shachnai\(^‡\)

May 26, 2022

Abstract

We study the 2-Dimensional Vector Bin Packing Problem (2VBP), a generalization of classic Bin Packing that is widely applicable in resource allocation and scheduling. In 2VBP we are given a set of items, where each item is associated with a two-dimensional volume vector. The objective is to partition the items into a minimal number of subsets (bins), such that the total volume of items in each subset is at most 1 in each dimension.

We give an asymptotic \((\frac{4}{3} + \varepsilon)\)-approximation for the problem, thus improving upon the best known asymptotic ratio of \((1 + \ln \frac{4}{3} + \varepsilon) \approx 1.406\) due to Bansal, Eliáš and Khan \(^2\). Our algorithm applies a novel Round\&Round approach which iteratively solves a configuration LP relaxation for the residual instance and samples a small number of configurations based on the solution for the configuration LP. For the analysis we derive an iteration-dependent upper bound on the solution size for the configuration LP, which holds with high probability. To facilitate the analysis, we introduce key structural properties of 2VBP instances, leveraging the recent fractional grouping technique of Fairstein et al. \(^12\).

\(^*\)CISPA Helmholtz Center for Information Security, Saarland Informatics Campus, Germany. ariel.kulik@cispa.de
\(^†\)Hamburg University of Technology, Institute for Algorithms and Complexity, Germany. matthias.mnich@tuhh.de
\(^‡\)Computer Science Department, Technion, Haifa, Israel. hadas@cs.technion.ac.il
1 Introduction

BIN PACKING (BP) is one of the most fundamental problems in combinatorial optimization. In BP we are given a set of items of different sizes that need to be packed in a minimum number of bins of unit capacities. The extensive study of Bin Packing since the early 1970’s has had a great impact on the design and analysis of approximation algorithms. In this work we study a generalization of BP, where the items to be packed as well as bin capacities are given as two-dimensional vectors.

The input for the 2-DIMENSIONAL VECTOR BIN PACKING PROBLEM (2VBP) is a pair \((I, v)\), where \(I\) is a set of \(n\) items and \(v : I \rightarrow [0, 1]^d\) is a two-dimensional volume function. A solution for the problem is a collection of subsets of items \(S_1, \ldots, S_m \subseteq I\) such that \(v(S_b) = \sum_{i \in S_b} v(i) \leq (1, 1)\) for all \(1 \leq b \leq m\) and \(\bigcup_{b=1}^m S_b = I\). The size of the solution is \(m\). Our objective is to find a solution of minimal size.

Apart from its theoretical significance, 2VBP is widely applicable in load balancing, cutting stock and multidimensional resource allocation in the cloud setting. A simple example is the allocation of computing services (items) to a minimal 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., [21, 23, 18, 27, 24].

Let \(OPT(I)\) be the value of an optimal solution for an instance \(I\) of a minimization problem \(P\) (see the standard definition, e.g., in [24]). As in the Bin Packing problem, we distinguish between absolute and asymptotic approximation. For \(\alpha \geq 1\), we say that \(A\) is an absolute \(\alpha\)-approximation algorithm for \(P\) if given an instance \(I\) of \(P\), \(A\) returns a solution of value at most \(\alpha \cdot OPT(I)\). \(A\) is an asymptotic \(\alpha\)-approximation algorithm for \(P\) if, for any instance \(I\), it returns a solution of value at most \(\alpha OPT(I) + o(OPT(I))\). If, for any fixed \(\varepsilon > 0\), there is an asymptotic \(\alpha\)-approximation algorithm where \(\alpha = 1 + \varepsilon\), then \(P\) admits an asymptotic polynomial-time approximation scheme (APTAS). We study randomized approximation algorithms for 2VBP. We say that \(A\) is a random asymptotic \(\alpha\)-approximation algorithm for a minimization problem \(P\) if, for any instance \(I\), \(A\) returns a solution of value at most \(\alpha OPT(I) + o(OPT(I))\) with constant probability.

In [19] Ray showed that there is no APTAS for 2VBP, assuming \(P \neq NP\). In the same work [19], Ray also showed the problem does not admit an asymptotic approximation ratio better than \(\frac{600}{399}\), assuming \(P \neq NP\). The best known asymptotic ratio for 2VBP prior to this work is \(1 + \ln \frac{2}{3} + \varepsilon \approx 1.406\) for any \(\varepsilon > 0\), due to Bansal, Eliáš and Khan [2]. Our main result is an improved asymptotic approximation ratio for the problem.

**Theorem 1.1.** For any \(\varepsilon > 0\) there is a polynomial-time random asymptotic \((\frac{4}{3} + \varepsilon)\)-approximation for 2VBP.

The bound in Theorem 1.1 is achieved by an algorithm that applies our novel Round&Round approach. The key idea is to iteratively obtain a fractional solution for a configuration LP relaxation of the (residual) problem, and use it to randomly sample a small part of the (integral) solution for the given instance. The analysis of the algorithm focuses on deriving an iteration-dependent bound on the size of an optimal solution for the configuration LP.

1.1 Related Work

The 2-Dimensional Vector Bin Packing Problem is commonly studied through the lens of the d-DIMENSIONAL VECTOR BIN PACKING PROBLEM (d-VBP), in which the volume of each item is a d-dimensional vector. Formally, for \(d \in \mathbb{N}_+\), the input for d-VBP is a pair \((I, v)\) where \(I\) is the set of items, and \(v : I \rightarrow [0, 1]^d\). As in 2VBP, a solution is a collection of subsets \(S_1, \ldots, S_m \subseteq I\) such that \(\bigcup_{b=1}^m S_b = I\) and \(v(S_b) = \sum_{i \in S_b} v(i) \leq 1\) for all \(1 \leq b \leq m\) and \(t \in \{1, \ldots, d\}\). The size of the solution is \(m\), and the objective is to find a solution of minimal size.

The one dimensional case (1-VBP) is the classic Bin Packing problem. A simple reduction from PARTITION shows there is no \(\alpha\)-approximation for Bin Packing with \(\alpha < \frac{2}{3}\), assuming \(P \neq NP\) (see, e.g., Chapter 9 in [25]). This motivates the study of asymptotic approximations for the problem. The

1\(^{\text{We say that } (a_1, a_2) < (b_1, b_2) \text{ if } a_1 < b_1 \text{ and } a_2 < b_2.}\)

2\(^{\text{Ray’s result addresses an oversight in an earlier proof of Woeginger [26].}}\)
first APTAS for Bin Packing was proposed by Fernandez de la Vega and Lueker [9], who introduced the linear grouping technique. The paper [15] gives an approximation algorithm that uses at most $OPT + O(\log^2(OPT))$ bins, where OPT is the size of the optimal solution for the instance. Recently, Hoberg and Rothvoß [14] obtained a polynomial-time algorithm that returns a solution of size $OPT + O(\log(OPT))$. For comprehensive surveys of known results for BP see, e.g., [7, 10].

For $d \geq 2$, an asymptotic approximation ratio arbitrarily close to $d$ can be easily obtained via a reduction to Bin Packing. The first non-trivial approximation for the bound for $OPT + \epsilon$ constraints ensure that each item is packed using the subset oblivious algorithm. The properties of the subset oblivious algorithm combined with a concentration bound of McDiarmid [17] are used to obtain the approximation guarantee. Round&Approx is the key framework used to obtain the best approximation algorithms for 2-D geometric Bin Packing and for Vector Bin Packing prior to this work.

The asymptotic $(1 + \ln \frac{d}{2} + \epsilon)$-approximation of [2] for 2VBPP is derived through an intricate combination of components in the Round&Approx framework with multibudgeted matching [11], along with reduction of the number of distinct volume vectors for the items. This reduction is achieved by applying linear grouping in one dimension and simple rounding of item volumes to multiples of a constant in the other dimension. Since rounding is performed over an optimal solution, the algorithm in practice guesses (enumerates) properties of this rounding. This results in a fairly complex algorithm. During the work on this paper we encountered a flaw in the analysis given in [2] We propose (in Section 4.3) a more sophisticated rounding technique that enables to resolve the issue for $d = 2$, albeit the fix is non-trivial.

Our rounding technique relies on a fractional version of the linear grouping technique applied in [2]. The fractional grouping technique was introduced in [12] for solving optimization problems under multiple knapsack constraints. The technique defines a partition of a set of items into groups based on a solution for a configuration LP relaxation of the problem at hand.

### 1.2 The Algorithm

Given a 2VBPP instance $(I, v)$, a configuration is a subset $C \subseteq I$ of items such that $v(C) \leq (1, 1) = 1$. Let $C(i) \in \{0, 1\}$ indicate whether item $i$ appears in $C$. We use $C$ to denote the set of all configurations. That is, $C = \{C \subseteq I \mid v(C) \leq (1, 1)\}$. We use a variant of the standard configuration LP, which given a demand $\tilde{d} \in [0, 1]^I$ of items is defined as follows:

$$\text{LP}(\tilde{d}) : \min \sum_{C \in C} \bar{x}_C \quad \forall i \in I : \sum_{C \in C} \bar{x}_C \cdot C(i) = \tilde{d}_i \quad \forall C \in C : \bar{x}_C \geq 0$$

Each of the variables $\bar{x}_C$ represents a (fractional) selection of the configuration $C$, where the first constraints ensure that each item $i \in I$ is covered according to its demand $\tilde{d}_i$. We use $OPT_f(\tilde{d})$ to denote the value of an optimal solution for LP(\tilde{d}). Our algorithms always use an integral demand

---

3Specifically, in page 1575 the inequality before the words “Thus by Lemma 6.1,” does not hold.
4We use the notation $1 = (1, \ldots, 1)$ and $0 = (0, \ldots, 0)$. 

---

2
vector $\bar{d} \in \{0, 1\}^L$, while the analysis considers fractional demand vectors. It is well known that there is a PTAS for LP$(\bar{d})$ with integral demand vectors.

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

We assume our instances adhere to a specific structure. Given $i \in I$ with high probability. We use a classification of items and configurations into categories. We say that $\bar{x}$ has a polynomial running time (we assume an algorithm) to generate a significant part of the solution following the random sampling stage.

This stands in contrast to algorithms associated with the framework (e.g., [1, 2]) which solve the configuration LP once and utilize a subset oblivious algorithm (or a variant of such algorithm) to generate a significant part of the solution following the random sampling stage.

The analysis of Algorithm 1 relies on an iteration dependent bound on $\OPT(I, v)$ which holds with high probability. We use a classification of items and configurations into categories. We say that an item $i \in I$ is $\delta$-large if $v_1(i) > \delta$ or $v_2(i) > \delta$ and use $L$ to denote the set of all $\delta$-large items ($\delta$ is commonly known by context). It can be easily shown that $|C \cap L| \leq 2 \cdot \delta^{-1}$ for all $C \in C$. For every $2 \leq h \leq 2 \cdot \delta^{-1}$ we define

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

Similarly, for a set of configurations $C' \subset C$, we use the indicator vector $1_{C'} \in \{0, 1\}^C$ in which entries corresponding to $C \in C'$ are equal to ‘1’.

Lemma 1.2. For any $\alpha \geq 1$ and $\delta \in (0, 0.1)$, if there is a random asymptotic $\alpha$-approximation for $\delta$-2VBP then there is a random asymptotic $(\alpha + 4\delta)$-approximation for 2VBP.

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

Our algorithm consists of two phases, the second of which is given in Algorithm 1. We note that Round&Round has a polynomial running time (we assume $\delta$ is fixed), and that it returns a solution for the 2VBP instance $(S_0, v)$. The distinction between $I$ and $S_0$ will be used later by our algorithm for 2VBP (see Algorithms 2). For a set of items $S \subseteq I$, we denote by $\mathbf{1}_S \in \{0, 1\}^I$ an indicator vector in which entries corresponding to $i \in S$ are equal to ‘1’, and all other entries are equal to ‘0’.

Algorithm 1: Round&Round

| Parameters: | $0 < \delta < 0.1$, $\alpha = -\ln (1 - \delta)$ and $k = \lceil \alpha \|\bar{x}\| \rceil$, where $\delta^{-1} \in \mathbb{N}$. |
| Input : | A $\delta$-2VBP instance $(I, v)$, $S_0 \subseteq I$ |
| 1 for | $j = 1, \ldots, k$ |
| 2 | Find a $(1 + \delta^j)$-approximate solution $\bar{x}^j$ for LP($\mathbf{1}_{S_{j-1}}$) and let $z_j = \|\bar{x}^j\|$ be its value. |
| 3 | Independently sample $\rho_j = [\alpha z_j]$ 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 $S_j \leftarrow S_{j-1} \setminus \bigcup_{\ell=1}^{\rho_j} C_{\ell}^j$. |
| 5 end |
| 6 Pack $S_k$ into configurations $C_1^*, \ldots, C_{\rho^*}^*$ using First-Fit |
| 7 Return $(\bigcup_{j=1}^{k} \{C_1^j, \ldots, C_{\rho_j}^j\}) \cup \{C_1^*, \ldots, C_{\rho^*}^*\}$. |

Step 4 of Algorithm 1 uses classic First-Fit to pack the remaining items (see Section 2 for more details).

In the analysis we show that $\rho^*$ is negligible in comparison to $\OPT(I, v)$. Thus, the solution generated by Algorithm 1 is comprised 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., [1, 2]) which solve the configuration LP once and utilize a subset oblivious algorithm (or a variant of such algorithm) to generate a significant part of the solution following the random sampling stage.

The analysis of Algorithm 1 relies on an iteration dependent bound on $\OPT_f(\mathbf{1}_{S_j})$ which holds with high probability. We use a classification of items and configurations into categories. We say that an item $i \in I$ is $\delta$-large if $v_1(i) > \delta$ or $v_2(i) > \delta$ and use $L$ to denote the set of all $\delta$-large items ($\delta$ is commonly known by context). It can be easily shown that $|C \cap L| \leq 2 \cdot \delta^{-1}$ for all $C \in C$. For every $2 \leq h \leq 2 \cdot \delta^{-1}$ we define

$$C_h = \{C \in C \mid v(C \cap L) > (1 - \delta, 1 - \delta) \text{ and } |C \cap L| = h\}. \quad (2)$$
Let $C_0 = C \backslash \left( \bigcup_{h=2}^{2\delta-1} C_h \right)$ be the set of all remaining configurations. As we assume that $(I, v)$ is an instance of $\delta$-2VBP (i.e., no $\delta$-huge items), it follows that for every $C \in C_0$ either $v_1(C \cap L) \leq 1 - \delta$ or $v_2(C \cap L) \leq 1 - \delta$.

For vectors $\bar{x}, \bar{z} \in [0, 1]^C$, $\bar{x} \cdot \bar{z} = \sum_{C \in \mathcal{C}} \bar{x}_C \cdot \bar{z}_C$ is the dot product of $\bar{x}$ and $\bar{z}$. It can be shown that if $\bar{x}^* \in [0, 1]^C$ is a solution for $\text{LP}(1_{S_0})$ then, ignoring negligible factors and with high probability, the solution returned by Algorithm 1 is of size at most

$$\bar{x}^* \cdot 1_{C_0} + \sum_{h=2}^{2\delta-1} \frac{h+1}{h} \cdot \bar{x}^* \cdot 1_{C_h} \leq \frac{3}{2} \|\bar{x}\|^2. \tag{3}$$

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 which is arbitrarily close to $3/2$. While we do not include a proof of (3), the proof can be derived by modifying the proof of Lemma 3.7 and using Lemma 3.10.

Our analysis relies on structural properties of 2VBP instances (inspired by properties presented in [2]) by which configurations in $C_0$ are “easy” (when selected by $\bar{x}^*$) and configuration in $C \backslash C_0$ are “difficult”. Intuitively, from the viewpoint of Round&Round, a configuration $C \in C \backslash C_0$ becomes easy at iteration $j$ if $C \cap L \subseteq S_j$, as in this case $C \cap S_j \in C_0$. Our analysis exploits this intuition via the notion of touched and untouched configurations (see the formal definition in Section 3.2). In contrast, the algorithm of [2] converts configurations in $C \backslash C_0$ to configurations in $C_0$ using a process which inflates the number of configurations in the solution. This inflation harms the approximation ratio (we elaborate on that in Section 3.3).

The bound in (3) suggests that the most “difficult” configurations in $\bar{x}^*$ are those in $C_2$; indeed, if we have an optimal solution containing no configuration in $C_2$ then we can obtain an approximation ratio of $4/3$. On the other hand, if an optimal (integral) solution contains only configurations in $C_2$ then a nearly optimal solution can be easily constructed using matching. As a solution may contain both configurations in $C_2$ and in $C \backslash C_0$, we use a sophisticated combination of a matching polytope and a configuration LP, along with the depended sampling technique of [4]. 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 $C_2$ and handled using matching techniques, and another which contains the remaining configurations that is handled by Round&Round.

We define the $\delta$-matching graph $G = (L, E)$ of $(I, v)$ as the graph in which the set of vertices consists of the $\delta$-large items, and $E = \{\{i_1, i_2\} \subseteq L \mid \{i_1, i_2\} \in C_2\}$. We use $P_M(G)$ to denote the matching polytope of $G$. We refer the reader to [22] for a formal definition of the matching polytope. Given $\bar{x} \in [0, 1]^C$, we define the projection of $\bar{x}$ on $E$ as the vector $\bar{p} \in \mathbb{R}_{E \geq 0}^L$ where $\bar{p}_e = \sum_{C \subseteq L \text{ s.t. } e \in C} \bar{x}_C$. 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 instance is the following optimization problem.

$$\text{MLP} : \quad \min \quad \sum_{C \in \mathcal{C}} \bar{x}_C$$

$$\forall i \in I : \quad \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = 1$$

$$\forall C \in \mathcal{C} : \quad \mathcal{E}(\bar{x}) \in P_M(G)$$

$$\forall C \in \mathcal{C} : \quad \bar{x}_C \geq 0 \tag{4}$$

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

Similar to the configuration LP, MLP can be approximated as well. The input for the MLP-problem is a $\delta$-2VBP instance $(I, v)$. A solution is a vector $\bar{x} \in \mathbb{R}_{C \geq 0}^C$ 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.
Lemma 1.3. For any $\delta \in (0, 0.1)$ there is a PTAS for the MLP-problem.

As $P_{M}(G)$ is defined using a non-polynomial number of constraints, both MLP and its dual have non-polynomial number of variables and a non-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 [10]), and more sophisticated tools are required to obtain a PTAS. The proof of Lemma 1.3 is given in Section 6.

Given $\tilde{x}$ such that $\tilde{\beta} = \mathcal{E}(\tilde{x}) \in P_{M}(G)$ and a parameter $\gamma > 0$, we use an algorithm of [3] to generate a random matching $\mathcal{M}$. Let $\mathcal{M} = \text{SampleMatching}(\tilde{\beta}, \gamma)$. We note that SampleMatching is a polynomial-time algorithm. The algorithm guarantees that $\Pr(e \in \mathcal{M}) = (1 - \gamma)\tilde{\beta}_e$ and provides Chernoff-like concentration bounds for $\mathcal{M}$ (see Lemma 6.18 for details).

The pseudocode of Match&Round, our algorithm for 2VBP, is given in Algorithm 2. 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: $\delta$-

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

Our main result follows from the next lemma.

Lemma 1.4. For any $\delta \in (0, 0.1)$ Algorithm 2 is a random asymptotic $(\frac{4}{3} + O(\delta))$-approximation for $\delta$-

Using Lemmas 1.4 and 1.2 we have the statement of Theorem 1.1. We use the standard notation of $\wedge$ for the minimum of two vectors element-wise. The analysis of Algorithm 2 relies on a partition of the solution $\tilde{x}^0$ obtained in Step 2 into its two “matching” and “fractional” components: $\tilde{x}^0 \wedge 1_{C_2}$ and $\tilde{x}^0 \wedge 1_{C \setminus C_2}$. We show that (w.h.p and ignoring negligible factors) $|\mathcal{M}| \leq \tilde{x}^0 \wedge 1_{C_2}$. Furthermore, we exploit the fact that $\tilde{x}^0 \wedge 1_{C \setminus C_2}$ does not select configurations from $C_2$ to show that the number of configurations returned by Round&Round (when invoked in Step 3 of Algorithm 2) is bounded by $\frac{4}{3} \cdot \tilde{x}^0 \wedge 1_{C \setminus C_2} + \frac{1}{3} \cdot \tilde{x}^0 \wedge 1_{C_2}$ (with constant probability and ignoring negligible factors).

1.3 Technical Contribution

Our Round&Round approach suggests a novel framework for solving covering problems. While the elegant Round&Round framework relies on solving an LP relaxation of a given instance once, and then sampling configurations until each item is covered with some constant probability, our framework solves the LP iteratively, using a relatively small number of samples of configurations in each iteration. The remaining (uncovered) items are then packed using a simple greedy algorithm (i.e., First-Fit). The crux of the analysis is to show the decrease in the optimal solution size for the remaining instance along this process. We show the usefulness of this approach for improving the bound of 2 for 2VBP.

To this end, we use in the analysis new ideas and techniques on which we elaborate below.

A main tool in the analysis of 2 is an implicit structural lemma which relies on the notion of slackness. A configuration $C$ has $\delta$-slack if $v_1(C) \leq 1 - \delta$, or $v_2(C) \leq 1 - \delta$. The authors of 2 consider a subset $S \subseteq I$ of items and a packing $A_1, \ldots, A_m$ of $S$ in $m$ bins, where $A_b$ has $\delta$-slack for all $1 \leq b \leq m$. They show that, with high probability, if a random set $Q \subseteq S$ has $\Pr(i \in Q) = \beta$ for all $i \in S$ and Chernoff-like concentration bounds, then $Q$ can be packed in $\approx \beta m$ bins.

---

6The idea to use matching algorithms is inspired by [2]. However, matching plays somewhat different roles in the two algorithms. In particular, the MLP is introduced in this work.

7To simplify the presentation we do not optimize the constants.

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

5
The analysis of [2] heavily relies on the property that any optimal solution for 2VBP, given by $A_1, \ldots, A_{OPT}$, can be converted to another solution $C_1, \ldots, C_k$, where each configuration either has $\delta$-slack or contains a tight pair of items.\footnote{A pair of items $i_1, i_2 \in L$ is tight if $v(\{i_1, i_2\}) > (1 - \delta, 1 - \delta).$} Subsequently, the structural property can be used with the subset of configurations among $C_1, \ldots, C_k$ which have $\delta$-slack. The authors show how to cleverly inflate such an optimal solution into one consisting of $k = 1.5OPT$ configurations.

The Round&Round approach allows us to reduce the overhead incurred by the inflation step. We say that a configuration $C \in C$ is touched in iteration $j$ of Round&Round if the total volume of uncovered large items in $C$ in some dimension is at most $1 - \delta$.\footnote{Recall that uncovered items are included in $S_j$.} Intuitively, for such configurations we can skip the inflation step used by [2]. This enables to tighten the bound on the number of bins used for the solution. This also reflects the core idea behind Round&Round, in which as the sampling process (Step 1 of Algorithm 1) progresses, the remaining instance “becomes easier”. Due to dependencies, implementing this approach requires an intricate calculation, in which we lower bound the probability that a configuration $C \in C$ becomes touched by some iteration. As a consequence, the analysis requires a variant of the structural lemma in which configurations are fractionally selected in accordance to the lower bound. Furthermore, whereas [2] used an integral solution ($A_1, \ldots, A_{OPT}$) as a reference point for the analysis, our analysis uses the non-matching part of $\overline{x}^0$ (that is, $\overline{x}^0 \land 1_{C(C_j)}$) as reference. As such, the limitations of the structural lemma of [2], allowing it to tackle integral solutions, render it too weak for our setting.

To overcome the above-mentioned difficulties we introduce a fractional version of the structural lemma of [2] (see Definition 2.1 and Lemma 2.2). Denote by $C^*$ the set of all multi-configurations, i.e., configurations $C$ which may contain multiple copies, $C(i) > 0$, of any item $i \in I$ (see the precise definition in Section 2). Let $\lambda \in [0, 1)^{C^*}$ be a vector such that any $C \in supp(\lambda)$ has $\delta$-slack and denote by $\overline{w}_i = \sum_{C \in C^*} \lambda_C \cdot C(i)$ the total (fractional) coverage of item $i$ by $\lambda$, for any $i \in I$. We show that if $d \in [0, 1]^I$ is a random demand vector such that $E[d_i] \leq \beta \overline{w}_i$, and $d$ satisfies some concentration bounds, then $OPT_\beta(d) \leq \beta \|\lambda\|$. To this end, we replace the classical linear grouping used in [2] (in one dimension of an item) by a fractional grouping technique introduced in [12] (see the details in Section 4.1).

1.4 Organization

In Section 2 we give some definitions and notation, as well as the key structural properties of 2VBP instances used in our analysis. Section 3 gives the analysis of Match&Round, along with a proof of its approximation guarantee, as stated in Lemma 1.4.

In Section 4 we prove Lemma 2.2 our main Structural Lemma. Section 5 gives the proofs of two structural properties used to obtain configurations with $\delta$-slack (stated as Lemmas 2.4 and 2.5). Finally, in Section 6 we show that MLP admits a PTAS (stated as Lemma 1.3).

2 Preliminaries

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

Recall that $OPT(I, v)$ is the minimal solution size for the instance $(I, v)$. Also, given $\overline{u} \in \mathbb{R}^I$ define the tolerance of $\overline{u}$ by $tol(\overline{u}) = \max \{ \sum_{i \in C} \overline{u}_i \mid C \in C^* \}$. Intuitively, the vector $\overline{u}$ associates with each item $i \in I$ some weight, $\overline{u}_i$; $tol(\overline{u})$ is the maximal total weight of any configuration $C$ w.r.t. $\overline{u}$. Our analysis relies on the existence of Linear Structures.
Definition 2.1 (Linear Structure). Let $\delta, K > 0$, $(I, v)$ be a $\delta$-2VBP instance, $\lambda \in [0,1]^{C'}$ and $\bar{w} \in [0,1]^I$ be the coverage of $\lambda$. A $(\delta, K)$-linear structure of $\lambda$ is $S \subseteq \mathbb{R}^I_{\geq 0}$ where $|S| \leq K$, and $S$ satisfies the following property. For any small items integral vector $\bar{z} \in [0,1]^I$ and $\beta \in [\delta^5, 1]$ such that $\text{supp}(\bar{z}) \subseteq \text{supp}(\bar{w})$ and

$$\forall \bar{u} \in S : \quad \bar{z} \cdot \bar{u} \leq \beta \cdot \bar{w} \cdot \bar{u} + \frac{1}{K^{10}} \cdot \text{OPT}(I, v) \cdot \text{tol}(\bar{u}),$$

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

Intuitively, a linear structure implies that if a demand vector $\bar{z}$ satisfies a ‘small’ number of constraints w.r.t. $\beta$ ($K$ is a constant, as defined in Lemma 2.2) then we obtain a decrease in $\text{OPT}_f(\bar{z})$ by factor of $\beta$. While linear structures do not necessarily exist for arbitrary vectors $\lambda$, we show that such structures exist for vectors which only select configurations with slack. We say that $C \in C^*$ has $\delta$-slack in dimension $d \in \{1, 2\}$ if $v_d(C) \leq 1 - \delta$. We say that $C \in C^*$ has $\delta$-slack if there is $d \in \{1, 2\}$ such that $C$ has $\delta$-slack in dimension $d$. Finally, we say $\lambda \in [0,1]^{C'}$ is with $\delta$-slack if for any $C \in \text{supp}(\lambda)$ it holds that $C$ has $\delta$-slack.

Lemma 2.2 (Structural Lemma). Let $(I, v)$ be a $\delta$-2VBP instance, where $\delta \in (0,0.1)$, and $\delta^{-1} \in \mathbb{N}$. There is a set $S^* \subseteq \mathbb{R}^I_{\geq 0}$ such that $|S^*| \leq \varphi(\delta) \cdot |I|^4$, where $\varphi(\delta) = \exp(\delta^{-20})$, which satisfies the following property.\(^{13}\) For any small items integral $\lambda \in [0,1]^{C'}$ with $\delta$-slack, there is a $(\delta, \varphi(\delta))$-linear structure $S$ of $\lambda$ where for all $\bar{u} \in S$: if $\text{supp}(\bar{u}) \cap L \neq \emptyset$ then $\bar{u} \in S^*$.

The proof of the lemma (given in Section 4) uses some of the structural properties shown in [2], along with the recent concept of fractional grouping, adopted from [12]. While the set $S^*$ 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 property is crucial for our analysis.

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

Definition 2.3. Given $C \in C$ and $\psi \geq 1$, we say that $\lambda \in [0,1]^{C'}$ is a $\psi$-relaxation of $C$ if the following conditions hold:

1. $\lambda$ is with $\delta$-slack.
2. $\|\lambda\| \leq \psi$.
3. $\sum_{C' \in C'} \lambda_{C'} \cdot C'(i) = C(i) \text{ for every } i \in I$.

Lemma 2.4. Let $\delta \in (0,0.1)$ s.t. $\delta^{-1} \in \mathbb{N}$, then for any $C \in C_0$ there is a $(1 + 4\delta)$-relaxation of $C$.

Lemma 2.5. Let $\delta \in (0,0.1)$. For any $2 \leq h \leq 2\delta^{-1}$ and $C \in C_h$ There is an $\frac{h}{h+1}$-relaxation of $C$.

Lemma 2.6. Let $\delta \in (0,0.1)$ and $C \in C$ such that $v(C) \leq (\delta, \delta)$, then there is a $4\delta$-relaxation of $C$.

The proofs of Lemmas 2.2, 2.4, and 2.5 are given in Section 5. Some of the statements and techniques used in the proofs can be viewed as variants of Lemma 5.3 in [2].

We use algorithm First-Fit in several places. The input for the algorithm is a 2VBP 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 BP (see, e.g., Chapter 9 of [25]), by taking for each item $i \in I$ in the 2VBP instance $\hat{v}_1(i) = \hat{v}_2(i) = \max\{v_1(i), v_2(i)\}$, and considering the problem in a single dimension.

Lemma 2.7. Given a 2VBP instance $(I, v)$ and a subset of items $S \subseteq I$, First-Fit returns a packing of $S$ in at most $2 \cdot (v_1(S) + v_2(S)) + 1$ bins.

\(^{13}\)Throughout the paper, for $x \in \mathbb{R}$, $\exp(x) = e^x$, where $e = 2.718\ldots$ is the base of the natural logarithm.
3 Analysis

In this section we give the analysis of Algorithm 2. Throughout this section we assume the \( \delta \)-2VBP instance \((I, v)\) and \( \delta \in (0, 0.1)\) are fixed. Thus, notations such as \( \rho_j, S_j, C^j_\ell \), and \( M \) refer to the corresponding variables in the execution of Algorithm 2 with \((I, v)\) as its input and \( \delta \) as the parameter. We also use \( \varphi(\delta) = \exp(\delta^{-20}) \) as in Lemma 2.2 and \( \text{OPT} = \text{OPT}(I, v) \). We commonly use \( k = \lfloor \ln_1(\delta) \rfloor \leq \delta^{-2} \).

In Section 3.1 we define the probabilistic space and prove some basic properties. The core of the analysis is in Section 3.2 which bounds the number of configuration sampled by \textit{Round&Round}. Section 3.3 gives the proof Lemma 1.4. 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 3.4.

3.1 Probability Space and Properties

We start with a formal definition of the probability space generated by the algorithm, denoted by \((\Omega, F, \text{Pr})\). Observe that \( \rho_j \leq [\alpha z_j] \leq \lfloor (\ln(1 - \delta)) (1 + \delta^2) \text{OPT} \rfloor \leq \text{OPT} \) for all \( j \in [k] \), as \( \delta < 0.1 \).\footnote{For any \( k \in \mathbb{R} \) we define \([k]\) = \{ \lfloor j \in \mathbb{N} \mid 1 \leq j \leq k \}.\} W.l.o.g assume that in each iteration Algorithm 1 samples \( \text{OPT} \) configurations \( C^j_1, \ldots, C^j_{\text{OPT}} \) independently according to \( \bar{x}^j \) and ignores configurations \( C^j_{\rho_j+1}, \ldots, C^j_{\text{OPT}} \). Furthermore, we may assume that \( \Omega \) is finite. Define the random variables \( P_0 = S_0 \) and \( P_j = (C^j_1, \ldots, C^j_{\text{OPT}}) \) for \( j \in [k] \).

Let \( F_j = \sigma(P_0, P_1, \ldots, P_j) \) be the \( \sigma \)-algebra of the random variables \( P_0, P_1, \ldots, P_j \). We also define \( F_{-1} = \{ \emptyset, \Omega \} \). It follows that \( F_{-1} \subseteq F_0 \subseteq F_1 \subseteq \ldots \subseteq F_k \).

We use conditional expectations and probabilities given the \( \sigma \)-algebra \( F_j \). We refer the reader to standard textbooks on probability (e.g., [2]) for the formal definitions. Intuitively, \( \mathbb{E}[X|F_j] \) 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 \( \alpha \) is set so the probability of \( i \in S_j \) decreases exponentially with \( j \), as shown in the next lemma. Given a boolean expression \( D \), we define \( \mathbb{I}_D \in \{0, 1\} \), where \( \mathbb{I}_D = 1 \) if \( D \) is true and \( \mathbb{I}_D = 0 \) otherwise. If \( D \) is random so is \( \mathbb{I}_D \).

**Lemma 3.1.** For any \( j \in [k] \) and \( i \in I \) it holds that \( \Pr(i \in S_j \mid F_{j-1}) \leq (1 - \delta) \cdot \mathbb{I}_{i \in S_{j-1}} \).

**Proof.** We can write

\[
\Pr(i \in S_j \mid F_{j-1}) = \mathbb{I}_{i \in S_{j-1}} \cdot \Pr(\forall \ell \in \rho_j : i \notin C^\ell_j \mid F_{j-1}) = \mathbb{I}_{i \in S_{j-1}} \prod_{\ell=1}^{\rho_j} \Pr(i \notin C^\ell_j \mid F_{j-1}) = \mathbb{I}_{i \in S_{j-1}} \left(1 - \frac{\mathbb{I}_{i \in S_{j-1}}}{z_j}\right)^{\rho_j} \leq \mathbb{I}_{i \in S_{j-1}} \cdot \exp(-\alpha) = \mathbb{I}_{i \in S_{j-1}} \cdot (1 - \delta).
\]

The first equality holds by the definition of \( S_j \), and the second holds since \( C^j_1, \ldots, C^j_{\rho_j} \) are conditionally independent given \( F_{j-1} \) (note that \( \rho_j \) is \( F_{j-1} \)-measurable). The third equality holds since \( \bar{x}^j \) is a solution to LP(1, \( S_{j-1} \)). The inequality uses \( \rho_j \geq \alpha z_j \) and \( (1 - \frac{1}{x})^x \leq \exp(-1) \) for \( x \geq 1 \). \( \square \)

Recall that \( \rho^* \) is the number of configurations used by First-Fit in Step 6 of Algorithm 1. By Lemma 3.1 it follows that \( \mathbb{E}[v_1(S_k) + v_2(S_k)] \leq (1 - \delta^k) (v_1(I) + v_2(I)) \leq 2 \cdot \delta \text{OPT} \), and by Lemma 2.7 we have \( \mathbb{E}[\rho^*] \leq 4\delta \text{OPT} + 1 \). The next lemma uses a concentration bound to show that, with high probability, \( \rho^* \) does not significantly deviate from its expectation.

**Lemma 3.2.** With probability at least \( 1 - \delta^{-2} \cdot \exp(-\delta^7 \cdot \text{OPT}) \) it holds that \( \rho^* \leq 16 \cdot \delta \cdot \text{OPT} + 1 \).

The proof of Lemma 3.2 is given in Section 3.2.

Observe that \( \mathbb{E}[|M|] = (1 - \delta^k) \cdot \bar{x}^0 \cdot \mathbb{I}_{C_2} \). We use the concentration bounds of [1] to show that w.h.p \( |M| \) is close to its expectation.

**Lemma 3.3.** It holds that \( |M| \leq \bar{x}^0 \cdot \mathbb{I}_{C_2} + \delta^2 \cdot \text{OPT} \) with probability at least \( 1 - \exp(-\delta^{10} \cdot \text{OPT}) \).
The proof of the lemma is given in Section 3.4.

The size of the solution returned by Algorithm 2 is $|M| + \sum_{j=1}^{k} \rho_j + \rho^*$. As Lemmas 3.2 and 3.3 give upper bounds for $|M|$ and $\rho^*$, it remains to derive a bound on $\sum_{j=1}^{k} \rho_j$, the total number of configurations sampled by Round&Round.

3.2 Analysis of Round&Round

Our analysis relies on the key notion of untouched configurations. We define the set of untouched configurations at iteration $j \in \{0, 1, \ldots, k\}$ by

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

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 $j \in [k]$. 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 any $j \in \{0, 1, \ldots, k\}$ and $0 \leq h \leq 2 \cdot \delta^{-1}$.

Intuitively, we view configurations in $\mathcal{C}_0$ as “easy” in comparison to configurations in $\mathcal{C} \setminus \mathcal{C}_0$ which are more difficult. This distinction stems from fact that we can only construct linear structure for configurations with slack (Lemma 2.2), and since we can attain slack for configuration in $\mathcal{C}_0$ with negligible overhead (Lemma 2.3). As such, the configurations in $U_j$ can be viewed as configurations which “remain difficult” after iteration $j$, and the configurations in $T_j$ are the configurations which “become easy” in iteration $j$.

Observe that

$$\sum_{j=1}^{k} \rho_j \leq k + \alpha(1 + \delta^2) \sum_{j=0}^{k-1} \text{OPT}_f(I_{S_j}) \leq k + (1 + 2\delta)\delta \sum_{j=0}^{k-1} \text{OPT}_f(I_{S_j}), \tag{6}$$

where the first inequality uses $\rho_j = [\alpha z_j] \leq \alpha(1 + \delta^2)\text{OPT}_f(I_{S_j}) + 1$, and the second inequality uses $\alpha(1 + \delta^2) \leq (1 + 2\delta)\delta$. In the following we derive an upper bound on $\delta \sum_{j=0}^{k-1} \text{OPT}_f(I_{S_j})$. By (6) this would imply a bound on $\sum_{j=1}^{k} \rho_j$, the number of configurations sampled by Round&Round.

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

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

Informally, $\bar{x}^*$ 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 $\mathcal{F}_{-1}$-measurable. It can be easily verified that $\bar{x}^* \cdot 1_{\mathcal{C}_h} = \bar{x}^0 \cdot 1_{U_{0,h}}$ for every $3 \leq h \leq 2 \cdot \delta^{-1}$ and $\bar{x}^* \cdot 1_{\mathcal{C}_0} = \bar{x}^* \cdot 1_{\mathcal{C}_0} = 0$. Furthermore, for any $C \in \text{supp}(\bar{x}^*)$ it holds that $C \subseteq S_0 \cap L$.

Let $\bar{y}^* \in [0, 1]^I$ be the coverage of $\bar{x}^*$. It follows that $\text{supp}(\bar{y}^*) \subseteq S_0 \cap L$. 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 $1_I - \bar{y}^*$. In the analysis we consider these coverage vectors separately, using the inequality

$$\delta \sum_{j=0}^{k-1} \text{OPT}_f(I_{S_j}) \leq \delta \sum_{j=0}^{k-1} \text{OPT}_f(I_{S_j} \land \bar{y}^*) + \delta \sum_{j=0}^{k-1} \text{OPT}_f(I_{S_j} \land (1_I - \bar{y}^*)) \tag{7}.$$  

The configurations in $\text{supp}(\bar{x}^*)$ are the configuration which remain “difficult” after the sampling of $\mathcal{M}$, and thus $\bar{y}^*$ represents the coverage of items by these difficult configurations. The remaining configurations are either in $T_0$ or in $\mathcal{C}_2$; as the configuration in $T_0$ are “easy”, we use them to compensate for items which were 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 3.1 and concentration bound.
Lemma 3.4. For $j \in \{0,1,\ldots,k\}$, let $\bar{\lambda} \in [0,1)^C$ be an $F_j$-measurable random vector, $\bar{w}$ be the coverage of $\bar{\lambda}$, $S$ be an $F_j$-measurable random $(\delta,\varphi(\delta))$-linear structure of $\bar{\lambda}$, and $\bar{d} \in [0,1]^I$ be a small items integral $F_j$-measurable random demand vector. Then,

$$\forall j \leq r \leq k : \quad \text{OPT}_f(\bar{d} \wedge 1_{S_j}) \leq (1 - \delta)^{r-j}(1 + 10\delta)\|\bar{\lambda}\| + \varphi(\delta) + \delta^{10}\text{OPT},$$

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

$$\xi = \Pr \left( \forall \bar{u} \in S : (\bar{1}_{S_j} \wedge \bar{d}) \cdot \bar{u} \leq \bar{w} \cdot \bar{u} + \frac{1}{\varphi(\delta)} \cdot \text{OPT} \cdot \text{tol}(\bar{u}) \right).$$

The proof of the lemma is given in Section 3.4. Note that by Lemma 3.1 we have that $\mathbb{E}[(\bar{1}_{S_j} \wedge \bar{d}) \cdot \bar{u} | F_j] \leq (1 - \delta)^{r-j} \cdot (\bar{1}_{S_j} \wedge \bar{d}) \cdot \bar{u}$ for any $\bar{u} \in S$. The proof of the lemma uses a concentration bound to show that with high probability $(\bar{1}_{S_j} \wedge \bar{d}) \cdot \bar{u} \lesssim (1 - \delta)^{r-j} \cdot (\bar{1}_{S_j} \wedge \bar{d}) \cdot \bar{u}$; the linear structure is used to bound $\text{OPT}_f(\bar{1}_{S_j} \wedge \bar{d})$, assuming the event in (8) occurs. We proceed to bound separately $\delta \sum_{j=0}^{k-1} \text{OPT}_f(\bar{1}_{S_j} \wedge \bar{y}^*)$ (see Lemma 3.7) and $\delta \sum_{j=0}^{k-1} \text{OPT}_f(\bar{1}_{S_j} \wedge (\bar{1}_I - \bar{y}^*))$ (see Lemma 3.10). The bound on $\delta \sum_{j=0}^{k-1} \text{OPT}_f(\bar{1}_{S_j} \wedge \bar{y}^*)$ is derived using the next lemmas.

Lemma 3.5. With probability at least $1 - \delta^{-10} \exp (-\delta^{50} \cdot \text{OPT})$ it holds that

$$\forall 2 \leq h \leq 2 \cdot \delta^{-1}, \ j \in [k] : \quad \mathbb{E} [\bar{x}^* \cdot 1_{T_{j,h}} | F_{j-1}] - \bar{x}^* \cdot 1_{T_{j,h}} \leq \delta^{20} \cdot \text{OPT}.$$ (9)

The proof, given in Section 3.3, is a simple application of a concentration bound.

Lemma 3.6. There exists $\mu : (0,0.1) \rightarrow \mathbb{R}_+$, independent of the instance $(I, v)$ and $\delta$, such that

$$\forall 2 \leq h \leq 2 \cdot \delta^{-1}, \ j \in [k] : \quad \bar{x}^* \cdot 1_{U_{j,h}} \geq (1 - \delta)^{h-j} \cdot \bar{x}^* \cdot 1_{U_{0,0}} - \delta^{10} \cdot \text{OPT} \text{ or } \text{OPT}_f(\bar{1}_{S_j}) \leq \mu(\delta),$$

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

The lemma follows from the inequality $\Pr (C \in U_{j,h} | F_{j-1}) \geq 1_{C \in U_{j-1,h}} \cdot \left( 1 - \frac{1}{z \bar{y}^*} \right)^{\alpha \bar{z}_{j+1}}$ which follows from Lemma 3.1. The observation that $\left( 1 - \frac{1}{z} \right)^{\alpha z_{j+1}} \rightarrow (1 - \delta)^h$ as $z \rightarrow \infty$, and Lemma 3.5. The dependence on $\mu$ in the lemma arises as the observation holds only if $z$ is sufficiently large. The proof is given in Section 3.4. Henceforth, we use $\mu$ to denote the function in Lemma 3.6.

Lemma 3.7. Assuming $\text{OPT} > \delta^{-30} \cdot \varphi(\delta) + \mu(\delta))$, with probability at least $1 - \varphi^4(\delta) \cdot \exp \left( -\frac{\text{OPT}}{\varphi^2(\delta)} \right)$ it holds that

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

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

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

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

$$\text{OPT}_f(\bar{y}^* \wedge 1_{S_r}) \leq \sum_{j=1}^{r} \text{OPT}_f(\bar{d}_j \wedge 1_{S_r}) + \text{OPT}_f(\bar{y}^* - \sum_{j=1}^{r} \bar{d}_j) \leq \sum_{j=1}^{r} \text{OPT}_f(\bar{d}_j \wedge 1_{S_r}) + \bar{x}^* \cdot 1_{U_r}. \quad (10)$$
We use Lemma 3.4 to bound the term $OPT_f (\tilde{d}^j \land \mathbb{1}_{S_j})$. To this end, we construct a vector $\tilde{\lambda} \in [0,1]^C$ for every $j \in [k]$. For any $2 \leq h \leq 2 \cdot \delta^{-1}$ and $C \in C_h$, let $\gamma_C$ be a $\frac{h}{11}$-relaxation of $C$. The existence of $\gamma_C$ is guaranteed by Lemma 2.3. We define

$$\forall j \in [k]: \quad \tilde{\lambda}^j = \sum_{C \in C \cap C_0} \tilde{x}_C \cdot (\Pr(C \in T_j \mid F_{j-1}) - \delta \cdot \mathbb{1}_{C \subseteq U_{j-1}}) \cdot \gamma_C,$$

and let $\tilde{w}^j$ be the coverage of $\tilde{\lambda}^j$. Since $U_{j-1}$ is $F_{j-1}$-measurable, it follows that $\tilde{\lambda}^j$ is $F_{j-1}$-measurable (and thus also $F_j$-measurable). Furthermore, since $\gamma_C$ is with $\delta$-slack for every $C \in C \setminus C_0$, it follows that $\tilde{\lambda}^j$ with $\delta$-slack for every $j \in [k]$.

Now, for every $i \in L$ and $j \in [k]$, we have

$$\mathbb{E} \left[ \tilde{d}_i^j \cdot \mathbb{1}_{i \in S_j} \mid F_{j-1} \right] = \mathbb{E} \left[ \sum_{C \in C} \mathbb{1}_{C \in T_j} \cdot \mathbb{1}_{i \in S_j} \cdot \tilde{x}_C \cdot C(i) \mid F_{j-1} \right]$$

$$= \mathbb{E} \left[ \sum_{C \in C \cap C_0} (\mathbb{1}_{C \in T_j} - \mathbb{1}_{C \in T_j} \cdot \mathbb{1}_{i \in S_j}) \cdot \tilde{x}_C \cdot C(i) \mid F_{j-1} \right]$$

$$= \sum_{C \in C \cap C_0} \left( \Pr(C \in T_j \mid F_{j-1}) - \mathbb{E} \left[ \mathbb{1}_{i \in S_j} \mathbb{1}_{C \subseteq U_{j-1}} \mid F_{j-1} \right] \right) \cdot \tilde{x}_C \cdot C(i).$$

The second equality uses $T_j \cap C_0 = \emptyset$ for $j \geq 1$, and the third equality uses $\mathbb{1}_{C \in T_j} \cdot \mathbb{1}_{i \in S_j} = \mathbb{1}_{C \subseteq U_{j-1}} \cdot \mathbb{1}_{i \in S_j}$ for any configuration $C$ such that $i \in C$. By Lemma 2.3.1, we have

$$\mathbb{E} \left[ \mathbb{1}_{i \in S_j} \mathbb{1}_{C \subseteq U_{j-1}} \mid F_{j-1} \right] \geq \delta \cdot \mathbb{1}_{C \subseteq U_{j-1}}$$

for any $C \in C$ and $i \in C \setminus L$. Furthermore, since $\gamma_C$ is a relaxation of $C$, we have that $C(i) = \sum_{C' \subseteq C} \gamma_C \cdot C'(i)$. By incorporating these into (12), we have

$$\mathbb{E} \left[ \tilde{d}_i^j \cdot \mathbb{1}_{i \in S_j} \mid F_{j-1} \right] \leq \sum_{C \in C \cap C_0} \tilde{x}_C \cdot (\Pr(C \in T_j \mid F_{j-1}) - \delta \cdot \mathbb{1}_{C \subseteq U_{j-1}}) \cdot \sum_{C' \subseteq C} \gamma_C \cdot C'(i)$$

$$= \sum_{C' \subseteq C} C'(i) \cdot \sum_{C \in C \cap C_0} \tilde{x}_C \cdot (\Pr(C \in T_j \mid F_{j-1}) - \delta \cdot \mathbb{1}_{C \subseteq U_{j-1}}) \cdot \gamma_C$$

for every $i \in L$ and $j \in [k]$. Also, for any $i \in I \setminus L$ and $j \in [k]$, it holds that $\mathbb{E} \left[ \tilde{d}_i^j \cdot \mathbb{1}_{i \in S_j} \mid F_{j-1} \right] = 0 \leq \tilde{w}_i^j$, as supp($\tilde{y}^j$) $\subseteq L$ and $\tilde{y}^j$ is the coverage of $\tilde{x}_i$.

By Lemma 2.2 there is a $(\delta, \varphi(\delta))$-linear structure $S_j$ of $\tilde{\lambda}$ for any $j \in [k]$.

Claim 3.8. For any $j \in [k]$ it holds that

$$\Pr \left( \forall \tilde{u} \in S_j : (\mathbb{1}_{S_j} \land \tilde{d}^j) \cdot \tilde{u} \leq \mathbb{E} \left[ (\mathbb{1}_{S_j} \land \tilde{d}^j) \cdot \tilde{u} \mid F_{j-1} \right] + \frac{OPT \varphi^{11}(\delta)}{\varphi^{11}(\delta)} \cdot \text{tol}(\tilde{u}) \right) \leq \exp \left( -\frac{OPT \varphi^{25}(\delta)}{\varphi^{25}(\delta)} \right).$$

The proof of Claim 3.8 (given in Section 3.3) follows from an application of a concentration bound. By (13), it holds that $\mathbb{E} \left[ \tilde{u} \cdot (\tilde{d}^j \land \mathbb{1}_{S_j}) \mid F_{j-1} \right] \leq \tilde{w}^j \cdot \tilde{u}$ for any $j \in [k]$ and $\tilde{u} \in S_j$; therefore,

$$\Pr \left( \forall \tilde{u} \in S_j : (\mathbb{1}_{S_j} \land \tilde{d}^j) \cdot \tilde{u} \leq \tilde{w}^j \cdot \tilde{u} + \frac{OPT \varphi^{11}(\delta)}{\varphi^{11}(\delta)} \cdot \text{tol}(\tilde{u}) \right)$$

$$\geq \Pr \left( \forall \tilde{u} \in S_j : (\mathbb{1}_{S_j} \land \tilde{d}^j) \cdot \tilde{u} \leq \mathbb{E} \left[ (\mathbb{1}_{S_j} \land \tilde{d}^j) \cdot \tilde{u} \mid F_{j-1} \right] + \frac{OPT \cdot \text{tol}(\tilde{u})}{\varphi^{11}(\delta)} \right) \geq \exp \left( -\frac{OPT \varphi^{25}(\delta)}{\varphi^{25}(\delta)} \right).$$

The last inequality is by Claim 3.8. Thus, by Lemma 2.3.4 with probability at least

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

it holds that

$$\forall j \in [k], \quad \forall r \leq k : \quad OPT_f (\tilde{d}^j \land \mathbb{1}_{S_r}) \leq (1 - \delta)^{r-j} (1 + 10\delta) \| \tilde{\lambda}^j \| + \varphi(\delta) + \delta^{10}OPT.$$
We henceforth assume that (14), as well as the statements of Lemmas 3.5 and 3.6 hold.

For every $j \in [k]$,

$$
\|\tilde{\lambda}^j\| = \sum_{C \in \mathcal{C} \cap \mathcal{C}_0} \bar{x}_C^* \cdot (\Pr(C \in T_j \mid \mathcal{F}_{j-1}) - \delta \cdot 1_{C \in U_{j-1}}) \cdot \|\tilde{\phi}_C^j\|
$$

$$
\leq \sum_{h=2}^{2^{-\delta_3-1}} \sum_{C \in \mathcal{C}_h} \bar{x}_C^* \cdot (\Pr(C \in T_j \mid \mathcal{F}_{j-1}) - \delta \cdot 1_{C \in U_{j-1}}) \cdot \frac{h}{h-1}
$$

$$
= \sum_{h=2}^{2^{-\delta_3-1}} \frac{h}{h-1} \left( \mathbb{E} \left[ \bar{x}_r^* \cdot 1_{T_{j,h}} \mid \mathcal{F}_{j-1} \right] - \delta \cdot \bar{x}_r^* \cdot 1_{U_{j-1,h}} \right)
$$

$$
\leq \sum_{h=2}^{2^{-\delta_3-1}} \frac{h}{h-1} \left( \bar{x}_r^* \cdot 1_{T_{j,h}} + \delta^{10} \cdot \text{OPT} - \delta \cdot \bar{x}_r^* \cdot 1_{U_{j-1,h}} \right)
$$

$$
\leq \sum_{h=2}^{2^{-\delta_3-1}} \frac{h}{h-1} \left( (1 - \delta) \bar{x}_r^* \cdot 1_{U_{j-1,h}} - \bar{x}_r^* \cdot 1_{U_{j,h}} \right) + \delta^{8} \cdot \text{OPT}.
$$

The first inequality holds since $\tilde{\phi}_C^j$ is a $\frac{h}{h-1}$-relaxation of $C$ for any $C \in \mathcal{C}_h$, the second inequality is by Lemma 3.5, and the last inequality uses $T_{j,h} = U_{j-1,h} \setminus U_{j,h}$.

Combining (14) and (15) with $\text{OPT} > \delta^{-30} \varphi(\delta)$, we have

$$
\frac{\text{OPT}_f(\bar{y}^* \wedge 1_{S_r})}{1 + 10\delta} \leq (1 - \delta)^{r-j} \sum_{h=2}^{2^{-\delta_3-1}} \frac{h}{h-1} \left( (1 - \delta) \bar{x}_r^* \cdot 1_{U_{j-1,h}} - \bar{x}_r^* \cdot 1_{U_{j,h}} \right) + \bar{x}_r^* \cdot 1_{U_r} + \delta^7 \text{OPT}
$$

for every $j \in [k]$ and $j \leq r \leq k$. Using the last inequality and (10), we obtain

$$
\frac{\text{OPT}_f(\bar{y}^* \wedge 1_{S_r})}{1 + 10\delta} \leq \sum_{j=1}^{r} (1 - \delta)^{r-j} \sum_{h=2}^{2^{-\delta_3-1}} \frac{h}{h-1} \left( (1 - \delta) \bar{x}_r^* \cdot 1_{U_{j-1,h}} - \bar{x}_r^* \cdot 1_{U_{j,h}} \right) + \bar{x}_r^* \cdot 1_{U_r} + \delta^5 \text{OPT}
$$

$$
= \sum_{h=2}^{2^{-\delta_3-1}} \frac{h}{h-1} \left( (1 - \delta)^r \cdot \bar{x}_r^* \cdot 1_{U_{0,h}} - \bar{x}_r^* \cdot 1_{U_{r,h}} \right) + \bar{x}_r^* \cdot 1_{U_r} + \delta^5 \text{OPT}
$$

$$
= \sum_{h=2}^{2^{-\delta_3-1}} \frac{1}{h-1} \left( (1 - \delta)^r \cdot \bar{x}_r^* \cdot 1_{U_{0,h}} - \bar{x}_r^* \cdot 1_{U_{r,h}} \right) + (1 - \delta)^r \bar{x}_r^* \cdot 1_{U_0} + \delta^5 \text{OPT}
$$

for every $r \in \{0, 1, \ldots, k \}$. Observe that $\text{OPT}_f(\bar{y}^* \wedge 1_{S_r}) \leq \text{OPT}_f(1_{S_r}) \leq \text{OPT}_f(1_{S_j})$ for any $1 \leq j \leq r$; thus, if $\text{OPT}(1_{S_j}) \leq \mu(\delta) \leq \delta^{30}\text{OPT}$ for some $j \in [k]$, then for every $r \geq j$ it holds that $\text{OPT}_f(\bar{y}^* \wedge 1_{S_r}) \leq \delta^{30}\text{OPT}$. Using the above inequality and Lemma 3.6, we have

$$
\frac{\text{OPT}_f(\bar{y}^* \wedge 1_{S_r})}{1 + 10\delta} \leq \sum_{h=2}^{2^{-\delta_3-1}} \frac{(1 - \delta)^r - (1 - \delta)^{h-r}}{h-1} \cdot \bar{x}_r^* \cdot 1_{U_{0,h}} + (1 - \delta)^r \bar{x}_r^* \cdot 1_{U_0} + \delta^{4}\text{OPT}.
$$
Thus,

\[
\delta \sum_{j=0}^{k-1} \text{OPT}_f (1_{S_j} \land \bar{y}^*) \geq \frac{\delta}{1 + 10\delta}
\]

\[
\leq \sum_{j=0}^{k-1} \sum_{h=2}^{2.5h-j} \frac{1}{h-1} \cdot \bar{x}^* \cdot 1_{U_0,h} + \delta \cdot \sum_{j=0}^{k-1} \frac{1}{h-1} \cdot (1 - \delta)^{h-j} \cdot 1_{U_0,h} + \delta^3 \text{OPT}
\]

\[
\leq \sum_{h=2}^{2.5h-j} \frac{1}{h-1} \left( \frac{1 - (1 - \delta)^h}{1 - (1 - \delta)^h} \right) + \delta \cdot \frac{1 - (1 - \delta)^h}{1 - (1 - \delta)} \cdot \bar{x}^* \cdot 1_{U_0} + \delta^3 \text{OPT}
\]

The second inequality holds since \((1 - \delta)^h \geq \delta\) and \(1 - (1 - \delta)^h \geq 1 - \delta h\). The last inequality uses \(\bar{x}^* \cdot 1_{U_0,h} = \bar{x}^0 \cdot 1_{U_0,h}\) for \(h \geq 3\), and \(\bar{x}^* \cdot 1_{C_2} = 0\) by the definition of \(\bar{x}^*\). Since \(\|\bar{x}^*\| \leq \|\bar{x}^0\| \leq (1 + \delta^3) \text{OPT} \leq 1.01 \cdot \text{OPT}\), we have

\[
\delta \sum_{j=0}^{k-1} \text{OPT}_f (1_{S_j} \land \bar{y}^*) \leq \sum_{h=3}^{2.5h-j} \frac{h+1}{h} \cdot \bar{x}^0 \cdot 1_{U_0,h} + 30 \cdot \delta \cdot \text{OPT} \leq \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{U_0 \setminus C_2} + 30 \cdot \delta \cdot \text{OPT},
\]

as in the statement of the lemma. As we assumed that \(|I|\) and the statements of Lemmas 3.5 and 3.6 hold, it follows the above inequality holds with probability at least

\[
1 - \varphi^3(\delta) \cdot \exp\left( - \frac{\text{OPT}}{\varphi^{25}(\delta)} \right) \geq 1 - \varphi^4(\delta) \cdot \exp\left( - \frac{\text{OPT}}{\varphi^{25}(\delta)} \right).
\]

Define \(\bar{y}^M\) as the coverage of \(\bar{x}^0 \land 1_{C_2}\); that is, \(\bar{y}^M = \sum_{C \in C_2} \bar{x}_C^0 \cdot C(i)\) for all \(i \in I\). To obtain a bound on \(\delta \sum_{j=0}^{k-1} \text{OPT}_f (1_{S_j} \land (1_I - \bar{y}^*))\), we use the next lemma.

**Lemma 3.9.** For any \(i \in I\) it holds that \(\Pr(i \notin S_0) = (1 - \delta^4) \bar{y}^M_i\) if \(i \in L\), and \(\Pr(i \notin S_0) = 0\) otherwise.

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

\[
\Pr(i \notin S_0) = \mathbb{E}[1_{i \notin S_0}] = \sum_{i' \in N(i)} \mathbb{E}[1_{(i, i') \in M}] = (1 - \delta^4) \sum_{i' \in N(i)} \sum_{C \in C_2 \text{ s.t. } (i, i') \subseteq C} \bar{x}_C^0
\]

\[
= (1 - \delta^4) \sum_{C \in C_2} \bar{x}_C^0 \cdot C(i) = (1 - \delta^4) \cdot \bar{y}^M_i.
\]

The third equality holds since \(\Pr(e \in M) = (1 - \delta^4) \sum_{C \in C_2 \text{ s.t. } e \subseteq C} \bar{x}_C^0\). Also, for any \(i \in I \setminus L\) it holds that \(i \notin \bigcup_{e \in M} e\); thus, \(i \in S_0\), i.e., \(\Pr(i \notin S_0) = 0\).

We now derive an upper bound for \(\delta \sum_{j=0}^{k-1} \text{OPT}_f (1_{S_j} \land (1_I - \bar{y}^*))\).

**Lemma 3.10.** Assuming \(\text{OPT} > \delta^{-30} \varphi(\delta)\), with probability at least \(1 - \exp\left( - \frac{\text{OPT}}{\varphi^{25}(\delta)} + \varphi^2(\delta) \cdot \ln \text{OPT} \right)\) it holds that

\[
\delta \sum_{j=0}^{k-1} \text{OPT}_f (1_{S_j} \land (1_I - \bar{y}^*)) \leq \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{U_0 \setminus C_2} + \frac{1}{3} \cdot |M| + 50 \delta (\text{OPT} + |M|).
\]
$$\lambda = \delta^4 \sum_{i \in L} \tilde{y}_i^C \cdot 1_{\{i\}} + \sum_{C \in T_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot \tilde{\gamma}_C + \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot \tilde{\tau}_C \setminus L,$$

where $1_{\{i\}} \in [0,1]^C = \tilde{z}$ such that $\tilde{z}(i) = 1$, and $\tilde{\gamma}_C = 0$ for $C \in C^* \setminus \{\{i\}\}$. Observe that $C_0 \subseteq T_0$ by definition; thus, $v(C \setminus L) \leq (\delta, \delta)$ for every $C \in U_0 \cup \mathcal{C}_2$. That is, $\lambda$ is well defined. Since the instance does not contain $\delta$-huge items it follows that $1_{\{i\}}$ is with $\delta$-slack. Hence, $\lambda$ is with $\delta$-slack as well. As $T_0$ and $U_0$ are $F_0$-measurable, it follows that $\lambda$ is $F_0$-measurable. Let $\tilde{w}$ be the coverage of $\lambda$ and define $\tilde{d} = 1_{S_0} \land (1 - \tilde{y}^*)$. Observe that we may have $\tilde{w}_i > 0$ (i.e., $i \in \text{supp}(\tilde{w})$) 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 $i \in L \setminus S_0$ for which $\tilde{w}_i < \tilde{d}_i$.

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

$$\tilde{w}_i = \sum_{C \in \mathcal{C}} \tilde{\lambda}_C \cdot C(i) = \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i) + \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i)$$

$$= \sum_{C \in \mathcal{C}} \tilde{x}_C^0 \cdot C(i) = 1 = 1_{i \in S_0}(1 - y_i^*) = \tilde{d}_i$$

The forth equality holds as $x^0$ is a solution for MLP. The fifth equality holds since $y_i^* = 0$ for all $i \in I \setminus L$ and by Lemma 3.9. In particular, it follows that $\tilde{w}$ and $\lambda$ are small items integral and $\tilde{w}_i - \tilde{d}_i = 0$ for any $i \in I \setminus L$.

For any $i \in L$, we have

$$\tilde{d}_i = 1_{i \in S_0} \left( 1 - \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i) \right)$$

$$= 1_{i \in S_0} - (1 - 1_{i \notin S_0}) \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i)$$

$$= 1_{i \in S_0} - \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i) - \sum_{C \in \mathcal{C}} 1_{i \notin S_0} \cdot 1_{C \in U_0} \cdot \tilde{x}_C^0 \cdot C(i)$$

$$= 1_{i \in S_0} - \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i),$$

where the forth 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 $i \in L$,

$$\tilde{w}_i - \tilde{d}_i = \delta^4 \cdot \tilde{y}_i^M + \sum_{C \in T_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i) - \left( 1_{i \in S_0} - \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i) \right)$$

$$= \delta^4 \cdot \tilde{y}_i^M + \sum_{C \in \mathcal{C}} \tilde{x}_C^0 \cdot C(i) - 1_{i \in S_0}$$

$$= \delta^4 \cdot \tilde{y}_i^M + 1 - y_i^M - 1_{i \in S_0}$$

$$= 1_{i \notin S_0} - (1 - \delta^4) \cdot \tilde{y}_i^M,$$

where the third equality holds since

$$1 = \sum_{C \in \mathcal{C}} \tilde{x}_C^0 \cdot C(i) = \sum_{C \in U_0 \cup \mathcal{C}_2} \tilde{x}_C^0 \cdot C(i) + \sum_{C \in \mathcal{C}} \tilde{x}_C^0 \cdot C(i) = \sum_{C \in \mathcal{C}} \tilde{x}_C^0 \cdot C(i) + \tilde{y}_i^M.$$
By (16), (17) and Lemma 3.9, it holds that $E[\bar{w}_i] = E[\bar{d}_i]$ for every $i \in I$.

Using the concentration bounds for SampleMatching, as given in [4], we can show that with high probability $\bar{d} \cdot \bar{u} \leq \bar{a} \cdot \bar{w}$ for every $\bar{u} \in \mathbb{R}^I_{\geq 0}$.

**Claim 3.11.** For any $\bar{u} \in \mathbb{R}^I_{\geq 0}$, it holds that

$$\Pr\left(\bar{d} \cdot \bar{u} > \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{21}(\delta)} \cdot \text{tol}(\bar{u})\right) \leq \exp\left(-\frac{\text{OPT}}{\varphi^{29}(\delta)}\right).$$

The proof of Claim 3.11 is given in Section 3.4.

Let $S^* \subseteq \mathbb{R}^I_{\geq 0}$ be the set defined in Lemma 2.2. Also, by Lemma 2.2, there exists a $(\delta, \varphi(\delta))$-linear structure $S$ of $\bar{\lambda}$ such that for any $\bar{u} \in S$ which satisfies supp$(\bar{u}) \cap L \neq \emptyset$ it holds that $\bar{u} \in S^*$. Observe that $S^*$ is non-random while $S$ is an $\mathcal{F}_0$-measurable random set, as $\bar{\lambda}$ is $\mathcal{F}_0$-measurable.

Claim 3.11 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 $\bar{u} \in S$. We use the set $S^*$ to circumvent this issue. Observe that for any $\bar{u} \in S$, if supp$(\bar{u}) \cap L = \emptyset$ then $\bar{d} \cdot \bar{u} = \bar{w} \cdot \bar{u}$ by (10), and if supp$(\bar{u}) \neq \emptyset$ then $\bar{u} \in S^*$. Thus,

$$\Pr\left(\forall \bar{u} \in S : \bar{d} \cdot \bar{u} \leq \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{21}(\delta)} \cdot \text{tol}(\bar{u})\right) \geq \Pr\left(\forall \bar{u} \in S^* : \bar{d} \cdot \bar{u} \leq \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{21}(\delta)} \cdot \text{tol}(\bar{u})\right)
\geq 1 - |S^*| \cdot \exp\left(-\frac{\text{OPT}}{\varphi^{29}(\delta)}\right) \geq 1 - \exp\left(-\frac{\text{OPT}}{\varphi^{29}(\delta)} + \varphi(\delta) \cdot \ln \text{OPT}\right).$$

The second inequality is by the union bound and Claim 3.11. The third inequality holds since $|S^*| \leq \varphi(\delta) \cdot |L|^4 \leq \varphi(\delta) \cdot 2^4 \cdot \delta^{-4} \cdot \text{OPT}^4$ as $\text{OPT} \geq \frac{d}{l}|L|$. Therefore, by Lemma 3.4, it holds that

$$\forall 0 \leq j \leq k : \text{OPT}_f(\mathbb{1}_{S_j} \land (\mathbb{1}_I - \bar{y}^*)) \leq (1 - \delta^3)\{(1 + 10\delta)\|\bar{\lambda}\| + \varphi(\delta) + \delta^{10}\text{OPT}\}$$

with probability at least

$$1 - \exp\left(-\frac{\text{OPT}}{\varphi^{29}(\delta)} + \varphi(\delta) \cdot \ln \text{OPT}\right) - \varphi^2(\delta) \cdot \exp\left(-\frac{\text{OPT}}{\varphi^{29}(\delta)}\right) \geq 1 - \exp\left(-\frac{\text{OPT}}{\varphi^{29}(\delta)} + \varphi(\delta) \cdot \ln \text{OPT}\right).$$

We henceforth assume that (18) holds.

We note that

$$\|\bar{\lambda}\| \leq \delta^4 \cdot \mathbb{1}_L \cdot \bar{y}^M + \sum_{h=3}^{2\delta^{-1}} \frac{h}{h-1} \cdot \mathbb{1}_L \cdot \bar{x}^0 \cdot \mathbb{1}_{T_{0,b}} + (1 + 4\delta) \cdot \mathbb{1}_C \cdot \mathbb{1}_0 \cdot 1 + 4\delta \|\bar{x}^0\|$$

$$\leq \frac{4}{3} \cdot \mathbb{1}_{T_{0,b}\setminus C_2} \cdot \bar{x}^0 \cdot \frac{1}{6} \cdot \sum_{h=3}^{2\delta^{-1}} \mathbb{1}_{T_{0,b}} + 10\delta \cdot \text{OPT},$$

where the second inequality uses

$$\mathbb{1}_L \cdot \bar{y}^M = \sum_{i \in L} \bar{y}^M_i = \sum_{i \in L} \sum_{C \in C_2} \bar{x}^0_i \cdot C(i) = \sum_{C \in C_2} \bar{x}^0 \cdot 2 \leq 2 \cdot \bar{x}^0 \cdot \mathbb{1}_C \leq 2 \cdot (1 + \delta^2)\text{OPT}.$$

It also holds that

$$\sum_{h=3}^{2\delta^{-1}} \bar{x}^0 \cdot \mathbb{1}_{T_{0,b}} = \sum_{C \in C_2} \bar{x}^0 \cdot \mathbb{1}_{C \setminus C_2} \leq \sum_{C \in C_2} \bar{x}^0 \sum_{i \in C \setminus L} \mathbb{1}_{i \not\in S_0} \leq \sum_{i \in L} \mathbb{1}_{i \not\in S_0} \sum_{C \in C_2} \bar{x}^0 \cdot C(i) \leq \sum_{i \in L} \mathbb{1}_{i \not\in S_0} \leq 2 \cdot |\mathcal{M}|.$$

Plugging the above inequality into (19), we obtain

$$\|\bar{\lambda}\| \leq \frac{4}{3} \cdot \bar{x}^0 \cdot \mathbb{1}_{T_{0,b}} + \frac{1}{3} \cdot |\mathcal{M}| + 10 \cdot \delta \cdot \text{OPT},$$

(20)
By (18) and (20), we have
\[
\delta \sum_{j=0}^{k-1} \OPT_f(1_{S_j} \cap (1 - \bar{y}^*)) \leq \delta \sum_{j=0}^{k-1} \left( (1 - \delta^2)(1 + 10\delta)\|\bar{x}\| + \varphi(\delta) + \delta^{10}\OPT \right)
\]
\[
\leq (1 + 10\delta)\|\bar{x}\| + \delta^8\OPT
\]
\[
\leq (1 + 10\delta) \left( \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{T_0 \backslash C_2} + \frac{1}{3} \cdot |M| + 10\delta \cdot \OPT \right) + \delta^8\OPT
\]
\[
\leq 4 \cdot \bar{x}^0 \cdot 1_{T_0 \backslash C_2} + \frac{1}{3} \cdot |M| + 50\delta(\OPT + |M|),
\]
where the second inequality uses \(\delta \geq 0.01 \cdot \OPT\). As we assumed (18) holds, the above inequality holds with probability at least \(1 - \exp \left(-\frac{\OPT}{\varphi^2(\delta)} + \varphi^2(\delta) \cdot \ln \OPT \right)\), as stated in the lemma.

### 3.3 Asymptotic Approximation Ratio

**Proof of Lemma 1.4.** Note that we may assume \(\OPT > \delta^{-30}(\varphi(\delta) + \mu(\delta))\). By Lemmas 3.7 and 3.10 we have
\[
\sum_{j=1}^{k} \rho_j \leq k + (1 + 2\delta)\delta \sum_{j=0}^{k-1} \OPT_f(1_{S_j})
\]
\[
\leq k + (1 + 2\delta) \left( \delta \sum_{j=0}^{k-1} \OPT_f(1_{S_j} \cap \bar{y}^*) + \delta \sum_{j=0}^{k-1} \OPT_f(1_{S_j} \cap (1 - \bar{y}^*)) \right)
\]
\[
\leq k + (1 + 2\delta) \left( \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{U_0 \backslash C_2} + 30\delta\OPT + \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{T_0 \backslash C_2} + \frac{1}{3} \cdot |M| + 50\delta(\OPT + |M|) \right)
\]
\[
\leq k + (1 + 2\delta) \left( \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{C \backslash C_2} + \frac{1}{3} \cdot |M| + 80\delta(\OPT + |M|) \right)
\]
\[
\leq 4 \cdot \bar{x}^0 \cdot 1_{C \backslash C_2} + \frac{1}{3} \cdot |M| + 90\delta(\OPT + |M|).
\]
The first inequality uses (18) and the last inequality assumes \(\OPT > \delta^{-30}(\varphi(\delta) + \mu(\delta))\). The number of configurations returned by the algorithm (assuming the statement of the lemmas hold) is
\[
|M| + \sum_{j=1}^{k} \rho_j + \rho^* \leq |M| + \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{C \backslash C_2} + \frac{1}{3} \cdot |M| + 90\delta(\OPT + |M|) + 16\delta\OPT + 1
\]
\[
\leq \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{C \backslash C_2} + 110\delta \cdot \OPT + \left( \frac{4}{3} + 90\delta \right) |M|
\]
\[
\leq \frac{4}{3} \cdot \bar{x}^0 \cdot 1_{C \backslash C_2} + 110\delta \cdot \OPT + \left( \frac{4}{3} + 90\delta \right) \cdot (\bar{x}^0 \cdot 1_{C_2} + \delta^2\OPT)
\]
\[
\leq \left( \frac{4}{3} + 90\delta \right) \|\bar{x}^0\| + 110\delta\OPT + 90\delta^3\OPT
\]
\[
\leq \left( \frac{4}{3} + 250\delta \right) \OPT
\]
where the last inequality uses \(\|\bar{x}^0\| \leq (1 + \delta^2)\OPT\).
3.4 Concentration

In this section we give the missing proofs of Sections 3.1 and 3.2.

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 \( \eta \)-bounded difference if for any \( \bar{x}, \bar{x}' \in A^m \) and \( r \in [m] \) such that \( \bar{x}_\ell = \bar{x}'_\ell \) for all \( \ell \in [m] \setminus \{r\} \) (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 following result is due to \[17\].

**Lemma 3.12** (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 \( \eta \)-bounded difference. Also, let \( X_1, \ldots, X_m \in A \) be independent random variables. Then for any \( t \geq 0 \),

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

To motivate our next lemma, consider the following example arising in our setting. Recall that \( T_{j,h} \) is the set of configurations in \( C_h \) that become touched in iteration \( j \), where \( j \in [k] \) and \( 2 \leq h \leq 2 \cdot \delta^{-1} \). For \( \bar{x} \in [0, 1]^C \), define the random variable \( 1_{T_{j,h}} \cdot \bar{x} \), which indicates the total (fractional) number of bins assigned to \( T_{j,h} \) according to \( \bar{x} \). We can write this random variable as a function of \( C_{j_1}^1, \ldots, C_{j_\ell}^1 \_OPT \).

For any \( U \subseteq C \) and \( \rho \in \{\text{OPT}\} \) define \( f_{U,\rho} : C_{\text{OPT}} \to \mathbb{R} \) by

\[
f_{U,\rho}(C_1, \ldots, C_{\text{OPT}}) = \sum_{C \in U} 1_{C \cap L \cap (U_{k \in [\rho]} C_t) \neq \emptyset} \cdot \bar{x}_C.
\]

Then it can be verified that \( 1_{T_{j,h}} \cdot \bar{x} = g(C_{j_1}^1, \ldots, C_{j_\ell}^1) \) where \( g = f_{U_{j-1},\rho_j} \). However, we cannot use Lemma 3.12 to show that \( 1_{T_{j,h}} \cdot \bar{x} \approx \mathbb{E}[1_{T_{j,h}} \cdot \bar{x}] \) with high probability, since the random variables \( C_{j_1}^1, \ldots, C_{j_\ell}^1_{\text{OPT}} \) are not independent, and the function \( g \) is random. Nonetheless, we note that at the end of iteration \( j-1 \) (Step 4 of Algorithm 1) the values of \( U_{j-1} \) and \( \rho_j \) are known (while \( \rho_j \) was not computed yet, its value does not depend on future random samples); thus, the function \( g = f_{U_{j-1},\rho_j} \) is known at iteration \( j \) of the algorithm. Furthermore, the random variables \( C_{j_1}^1, \ldots, C_{j_\ell}^1_{\text{OPT}} \) are independent (by definition) assuming we have the random samples of the first \((j-1)\) iterations. Therefore, we expect Lemma 3.12 to hold in this setting. More formally, since \( C_{j_1}^1, \ldots, C_{j_\ell}^1_{\text{OPT}} \) are conditionally independent given \( \mathcal{F}_{j-1} \), and \( g \) is a random function that is \( \mathcal{F}_{j-1} \)-measurable, we expect that \( g(C_{j_1}^1, \ldots, C_{j_\ell}^1_{\text{OPT}}) \approx \mathbb{E}[g(C_{j_1}^1, \ldots, C_{j_\ell}^1_{\text{OPT}})|\mathcal{F}_{j-1}] \). This is formalized in the next lemma.

**Lemma 3.13** (Generalized McDiarmid). Given a finite arbitrary set \( A \), \( m \in \mathbb{N}_+ \) and \( \eta > 0 \), let \( D \) be a finite family of \( \eta \)-bounded difference functions from \( A^m \) to \( \mathbb{R} \). Let \( (\Omega, \mathcal{F}, \Pr) \) be a probability space for which \( \Omega \) is finite, \( \mathcal{G} \subseteq \mathcal{F} \) a \( \sigma \)-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 \( t \geq 0 \),

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

Lemma 3.13 can be derived from Lemma 3.12 using standard arguments from probability theory (we omit the details). We use Lemma 3.13 to prove the following technical result.

**Lemma 3.14**. Let \( j \in \{0, 1, \ldots, k - 1\} \) and \( t > 0 \). Also, let \( \bar{u} \in \mathbb{R}_{\geq 0}^I \) be an \( \mathcal{F}_j \)-measurable random vector. Then,

\[
\Pr ( \bar{u} \cdot 1_{S_{j+1}} - (1 - \delta)\bar{u} \cdot 1_{S_j} > t \cdot \text{tol}(\bar{u}) ) \leq \exp \left( - \frac{2 \cdot t^2}{\text{OPT}} \right).
\]

**Proof.** Let \( A \) be the set of all possible values the random vector \( \bar{u} \) can take (formally, \( A = \{\bar{u}(\omega) | \omega \in \Omega\} \)). Since \( \Omega \) is finite it holds that \( A \) is also finite.

For any \( S \subseteq I, \rho \in \{\text{OPT}\} \) and \( \bar{a} \in A \) define \( f_{S,\rho,a} : \mathcal{C}_{\text{OPT}} \to \mathbb{R} \) by

\[
f_{S,\rho,a}(C_1, \ldots, C_{\text{OPT}}) = \begin{cases} \frac{\text{tol}(\bar{a})}{\text{tol}(\bar{a})} \cdot \bar{a} \cdot 1_{S \setminus (U_{t=1} C_t)} & \text{if } \text{tol}(\bar{a}) \neq 0 \\ 0 & \text{otherwise} \end{cases}
\]

See, e.g., in [5] for a formal definition of conditional independence.
Also, define \( D = \{ f_{S,\rho,\bar{a}} \mid S \subseteq I, \rho \in [\text{OPT}], \bar{a} \in A \} \). It can be easily verified that \( D \) is finite.

Let \( f_{S,\rho,\bar{a}} \in D \), \((C_1, \ldots, C_{\text{OPT}}), (C_1', \ldots, C'_{\text{OPT}}) \in C_{\text{OPT}} \) and \( r \in [\text{OPT}] \) such that \( C_\ell = C'_\ell \) for every \( \ell \in [\text{OPT}] \setminus \{r\} \). If \( \text{tol}(\bar{a}) = 0 \) or \( r > \rho \) then \( |f_{S,\rho,\bar{a}}(C_1, \ldots, C_{\text{OPT}}) - f_{S,\rho,\bar{a}}(C_1', \ldots, C'_{\text{OPT}})| = 0 \). Otherwise, let \( T = \bigcup_{\ell \in [\rho]\setminus\{r\}} C_\ell = \bigcup_{\ell \in [\rho]\setminus\{r\}} C'_\ell \). Then

\[
\left| f_{S,\rho,\bar{a}}(C_1, \ldots, C_{\text{OPT}}) - f_{S,\rho,\bar{a}}(C_1', \ldots, C'_{\text{OPT}}) \right| = \left| \frac{1}{\text{tol}(\bar{a})} \cdot \bar{a} \left( 1_{S \setminus T \cap C_r} - 1_{S \setminus T \cap C'_r} \right) \right|
\]

\[
= \frac{1}{\text{tol}(\bar{a})} \cdot \max_{i \in (S \cap C'_r) \setminus (C_r \cup T)} a_i - \sum_{i \in (S \cap C_r) \setminus (C'_r \cup T)} a_i
\]

\[
\leq \frac{1}{\text{tol}(\bar{a})} \cdot \text{tol}(\bar{a}) \leq 1.
\]

The second equality holds as \( S \setminus T \setminus C_r \setminus (S \setminus T \setminus C'_r) = (S \cup C'_r) \setminus (C_r \cup T) \) and symmetrically \( S \setminus T \setminus C'_r \setminus (S \setminus T \setminus C_r) = (S \cup C_r) \setminus (C'_r \cup T) \). 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 \( S_j, \rho_{j+1} \) and \( \bar{u} \) are \( \mathcal{F}_j \)-measurable it follows that \( g \) is \( \mathcal{F}_j \)-measurable. By the definition of \( g \) we have

\[
\text{tol}(\bar{u}) \cdot g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) = \bar{u} \cdot 1_{S_j \setminus \bigcup_{k=1}^{j+1} C_k^{j+1}} = \bar{u} \cdot 1_{S_{j+1}}.
\]

Furthermore,

\[
\mathbb{E}[\text{tol}(\bar{a}) \cdot g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) \mid \mathcal{F}_j] = \mathbb{E}[\bar{u} \cdot 1_{S_{j+1}} \mid \mathcal{F}_j]
\]

\[
= \sum_{i \in I} \bar{u}_i \cdot \Pr(i \in S_{j+1} \mid \mathcal{F}_j) \leq (1 - \delta) \sum_{i \in I} \bar{u}_i \cdot 1_{i \in S_j} = (1 - \delta) \cdot \bar{a} \cdot 1_{S_j},
\]

where the inequality is by Lemma \[\text{3.1}\]. Therefore,

\[
\Pr \left( (\bar{u} \cdot 1_{S_{j+1}} - (1 - \delta)\bar{u} \cdot 1_{S_j} > t \cdot \text{tol}(\bar{u})) \right)
\]

\[
\leq \Pr \left( (g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) - \mathbb{E}[g(C_1^{j+1}, \ldots, C_{\text{OPT}}^{j+1}) \mid \mathcal{F}_j]) > t \right)
\]

\[
\leq \exp \left( -\frac{2 \cdot t^2}{\text{OPT}} \right),
\]

where the last inequality is by Lemma \[\text{3.13}\].

We use Lemma \[\text{3.14}\] to show the following.

**Lemma 3.15.** 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,

\[
\Pr \left( \exists r \in [j, k] \cap \mathbb{N} : \bar{u} \cdot 1_{S_r} - (1 - \delta)^{r-j} \cdot \bar{u} \cdot 1_{S_j} > t \cdot \text{tol}(\bar{u}) \right) \leq \delta^{-2} \cdot \exp \left( -\frac{2 \cdot \delta^4 \cdot t^2}{\text{OPT}} \right).
\]
Proof. We note that

\[ \Pr \left( \exists r \in [j,k] \cap \mathbb{N} : \bar{u} \cdot 1_{S_r} - (1 - \delta)^{r-j} \cdot \bar{u} \cdot 1_{S_j} > t \cdot \text{tol}(\bar{u}) \right) \]

\[ = \Pr \left( \exists r \in [j,k] \cap \mathbb{N} : \sum_{\ell=j+1}^r \left( \bar{u} \cdot 1_{S_{\ell-1}} - (1 - \delta) \cdot \bar{u} \cdot 1_{S_{\ell-1}} \right) \cdot (1 - \delta)^{r-\ell} > t \cdot \text{tol}(\bar{u}) \right) \]

\[ \leq \Pr \left( \exists r \in [j+1,k] \cap \mathbb{N}, \ell \in [j+1,r] \cap \mathbb{N} : \left( \bar{u} \cdot 1_{S_{\ell}} - (1 - \delta) \cdot \bar{u} \cdot 1_{S_{\ell-1}} \right) \cdot (1 - \delta)^{r-\ell} > \frac{t}{r-j} \cdot \text{tol}(\bar{u}) \right) \]

\[ \leq \sum_{\ell=j+1}^r \Pr \left( \bar{u} \cdot 1_{S_{\ell}} - (1 - \delta) \cdot \bar{u} \cdot 1_{S_{\ell-1}} > \frac{t}{k} \cdot \text{tol}(\bar{u}) \right) \]

\[ \leq k \cdot \exp \left( -\frac{2 \cdot \left( \frac{t}{k} \right)^2}{\text{OPT}} \right) \]

\[ \leq \delta^{-2} \cdot \exp \left( -\frac{2 \cdot \delta^4 \cdot t^2}{\text{OPT}} \right) . \]

The first inequality holds since if a sum of \( n \) variables is greater than \( T \) there must be a variable with value greater than \( \frac{T}{n} \). The forth inequality is by Lemma 3.14 and the last inequality uses \( k \leq \delta^{-2} \).

We can now proceed to the proofs of Lemmas 3.2 and 3.4.

Proof of Lemma 3.2. Define \( \bar{u} \in [0,1]^I \) by \( \bar{u}_i = v_1(i) + v_2(i) \). For any \( C \in C \) it holds that \( \sum_{i \in C} \bar{u}_i = v_1(C) + v_2(C) \leq 2 \), therefore \( \text{tol}(\bar{u}) \leq 2 \). Furthermore, there is partition \( Q_1, \ldots, Q_{\text{OPT}} \) of \( I \) such that \( Q_\ell \) is a configuration for each \( \ell \in [\text{OPT}] \), therefore

\[ \bar{u} \cdot 1_{S_0} \leq \bar{u} \cdot 1_I = \sum_{\ell=1}^{\text{OPT}} \bar{u} \cdot 1_{Q_\ell} \leq \text{OPT} \cdot \text{tol}(\bar{u}) \leq 2 \cdot \text{OPT} . \]  

(21)

Recall that \( \rho^* \) is the number of configuration used by First-Fit in Step 6 of Algorithm 1. Using Lemma 2.7 we have

\[ \Pr(\rho^* > 16 \cdot \delta \cdot \text{OPT} + 1) \leq \Pr(\sum_{\ell=1}^{\text{OPT}} v_1(S_\ell) + v_2(S_\ell) > 8 \cdot \delta \cdot \text{OPT}) \]

\[ \leq \Pr(\bar{u} \cdot 1_{S_k} > 8 \cdot \delta \cdot \text{OPT}) \]

\[ \leq \Pr \left( \bar{u} \cdot 1_{S_k} - (1 - \delta)^k \cdot \bar{u} \cdot 1_{S_0} > 6 \cdot \delta \cdot \text{OPT} \right) \]

\[ \leq \Pr \left( \exists r \in [0,k] \cap \mathbb{N} : \bar{u} \cdot 1_{S_{r}} - (1 - \delta)^r \cdot \bar{u} \cdot 1_{S_0} > \text{tol}(\bar{u}) \cdot \delta \cdot \text{OPT} \right) \]

\[ \leq \delta^{-2} \cdot \exp \left( -\frac{2 \cdot \delta^4 \cdot t^2 \cdot \text{OPT}^2}{\text{OPT}} \right) \]

\[ \leq \delta^{-2} \cdot \exp (-\delta^7 \cdot \text{OPT}) . \]

The third inequality uses (21) and \( (1 - \delta)^k \leq \delta \). The fifth inequality is by Lemma 3.15. Hence, \( \Pr(\rho^* \leq 16 \cdot \delta \cdot \text{OPT} + 1) \geq 1 - \delta^{-2} \cdot \exp (-\delta^7 \cdot \text{OPT}) \).

Proof of Lemma 3.4. Let \( S = \{ \bar{u}^1, \ldots, \bar{u}^{\lceil \varphi(\delta) \rceil} \} \), where \( \bar{u}^\ell \) is an \( F_j \)-measurable random vector for every \( 1 \leq \ell \leq \lceil \varphi(\delta) \rceil \) (in case \( |S| < \lceil \varphi(\delta) \rceil \) the same vector may appear several times in \( \bar{u}^1, \ldots, \bar{u}^{\lceil \varphi(\delta) \rceil} \)).
As $S$ is a $(\delta, \varphi(\delta))$ linear structure it holds that

\[
\Pr \left( \forall r \in [j, k] \cap N : \text{OPT} \left( \bar{d} \land \mathds{1}_S \right) \leq (1 - \delta)^{-j} (1 + 10\delta) \| \lambda \| + \varphi(\delta) + \delta^{10} \cdot \text{OPT} \right)
\geq \Pr \left( \forall r \in [j, k] \cap N, \ell \in [\varphi(\delta)] : (\mathds{1}_S \land \bar{d}) \cdot \bar{u}^\ell \leq (1 - \delta)^{-j} \cdot \bar{w} \cdot \bar{u}^\ell + \frac{\text{OPT} \varphi^{10}(\delta)}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}^\ell) \right)
\geq \Pr \left( \forall \ell \in [\varphi(\delta)] : (\mathds{1}_S \land \bar{d}) \cdot \bar{u}^\ell \leq (1 - \delta)^{-j} \cdot \bar{w} \cdot \bar{u}^\ell \right)
\geq \frac{\xi - \sum_{i=1}^{\lfloor \varphi(\delta) \rfloor} \Pr \left( \exists r \in [j, k] \cap N : (\mathds{1}_S \land \bar{d}) \cdot \bar{u}^\ell > (1 - \delta)^{-j} \cdot (\mathds{1}_S \land \bar{d}) \cdot \bar{u}^\ell + \frac{\text{OPT} \varphi^{10}(\delta)}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}^\ell) \right)}{\Pr \left( \exists r \in [j, k] \cap N : (\mathds{1}_S \land \bar{d}) \cdot \bar{u}^\ell > (1 - \delta)^{-j} \cdot (\mathds{1}_S \land \bar{d}) \cdot \bar{u}^\ell + \frac{\text{OPT} \varphi^{10}(\delta)}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}^\ell) \right)}.
\]

The fourth equality follows from the union bound and the definition of $\xi$ in $\mathcal{S}$. The fifth inequality is by Lemma 3.15.

The following technical lemma will be used to prove Lemma 3.5.

**Lemma 3.16.** Let $j \in [k]$ and $2 \leq h \leq 2 \cdot \delta^{-1}$. Then

\[
\Pr \left( \mathbb{E} [\bar{x}^* \cdot \mathds{1}_{T_{j,h}} | \mathcal{F}_{j-1}] - \bar{x}^* \cdot \mathds{1}_{T_{j,h}} > \delta^{20} \cdot \text{OPT} \right) \leq 2 \cdot \exp (\delta^{27} \cdot \text{OPT}).
\]

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

For any $U \subseteq C$, $\rho \in \text{[OPT]}$ and $\bar{x} \in \mathcal{V}$ define $f_{U, \rho, \bar{x}} : C^{\text{OPT}} \rightarrow \mathbb{R}$ by

\[
f_{U, \rho, \bar{x}} (C_1, \ldots, C_{\text{OPT}}) = \bar{x} \cdot \mathds{1}_{C \subseteq U} \cap \mathcal{C}_n \left( \bigcup_{i=1}^r C_i \right) \land \mathcal{L} \neq \emptyset \cdot \sum_{C \subseteq U} \bar{x}_C \cdot \mathds{1}_C \cap \mathcal{C}_n \left( \bigcup_{i=1}^r C_i \right) \land \mathcal{L} \neq \emptyset.
\]

Define $D = \{ f_{U, \rho, \bar{x}} | U \subseteq C, \rho \in \text{[OPT]}, \bar{x} \in \mathcal{V} \}$. It follows that $D$ is a finite set.

Let $f_{U, \rho, \bar{x}} \in D$, $(C_1, \ldots, C_{\text{OPT}}), (C'_1, \ldots, C'_{\text{OPT}}) \in C^{\text{OPT}}$, and $r \in \text{[OPT]}$ such that $C_\ell = C'_\ell$ for every $\ell \in \text{[OPT]} \setminus \{ r \}$. If $r > \rho$ then $| f_{U, \rho, \bar{x}} (C_1, \ldots, C_{\text{OPT}}) - f_{U, \rho, \bar{x}} (C'_1, \ldots, C'_{\text{OPT}}) | = 0$. Otherwise, let $T = \bigcup_{i \in [\rho] \setminus \{ r \}} C_i$. It holds that

\[
| f_{U, \rho, \bar{x}} (C_1, \ldots, C_{\text{OPT}}) - f_{U, \rho, \bar{x}} (C'_1, \ldots, C'_{\text{OPT}}) | = \left| \sum_{C \subseteq U} \bar{x}_C \cdot \mathds{1}_C \cap \mathcal{C}_n \left( \bigcup_{i=1}^r C_i \right) \land \mathcal{L} \neq \emptyset \cdot \mathds{1}_C \cap \mathcal{C}_n \left( \bigcup_{i=1}^r C_i \right) \land \mathcal{L} \neq \emptyset \right| 
\leq \max \left\{ \sum_{C \subseteq \mathcal{C}} \bar{x}_C \cdot \mathds{1}_C \cap \mathcal{C} \land \mathcal{L} \neq \emptyset \cdot \mathds{1}_C \cap \mathcal{C} \land \mathcal{L} \neq \emptyset, \sum_{C \subseteq \mathcal{C}} \bar{x}_C \cdot \mathds{1}_C \cap \mathcal{C} \land \mathcal{L} \neq \emptyset \cdot \mathds{1}_C \cap \mathcal{C} \land \mathcal{L} \neq \emptyset \right\}.
\]

Furthermore,

\[
\sum_{C \subseteq \mathcal{C}} \bar{x}_C \cdot \mathds{1}_C \cap \mathcal{L} \neq \emptyset \leq \sum_{C \subseteq \mathcal{C}} \bar{x}_C \sum_{i \in \mathcal{C} \cap \mathcal{L}} C(i) = \sum_{i \in \mathcal{C} \cap \mathcal{L}} \sum_{C \subseteq \mathcal{C}} \bar{x}_C \cdot C(i) \leq | \mathcal{C} \cap \mathcal{L} | \leq 2 \cdot \delta^{-1},
\]
and by a symmetric argument \( \sum_{C \in \mathcal{E}} \mathbb{E}[C \cdot 1_{C \cap \mathcal{L} \neq \emptyset}] \leq 2 \cdot \delta^{-1} \). Thus,

\[
\left| f_{U,p,\tilde{x}}(C_1, \ldots, C_{\text{OPT}}) - f_{U,p,\tilde{x}}(C_1', \ldots, C_{\text{OPT}}') \right| \leq 2 \cdot \delta^{-1}.
\]

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

Define \( g = f_{U_j-1,h,\rho_j,\tilde{x}}, \) Since \( U_j-1,h, \rho_j \) and \( \tilde{x} \) 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 \( C \cap \mathcal{L} \cap \left( \bigcup_{\ell \in [\rho_j]} C_{\ell}^j \right) \neq \emptyset \). Thus,\( g(C_1^j, \ldots, C_{\text{OPT}}^j) = \tilde{x}^* \cdot 1_{\{C \in U_j-1,h \mid C \cap \left( \bigcup_{\ell \in [\rho_j]} C_{\ell}^j \right) \cap \mathcal{L} \neq \emptyset \}} = \tilde{x}^* \cdot 1_{T_{j,h}} \). Therefore,

\[
\Pr \left( \left| \mathbb{E} \left[ \tilde{x}^* \cdot 1_{T_{j,h}} \mid \mathcal{F}_{j-1} \right] - \tilde{x}^* \cdot 1_{T_{j,h}} \right| > 2 \cdot \delta^{20} \cdot \text{OPT} \right)
\]

The proof of Lemma 3.5 follows directly from Lemma 3.16.

**Proof of Lemma 3.6.** By the union bound, we have

\[
\Pr \left( \forall j \in [k], 2 \leq h \leq 2 \cdot \delta^{-1} : \left| \mathbb{E} \left[ \tilde{x}^* \cdot 1_{T_{j,h}} \mid \mathcal{F}_{j-1} \right] - \tilde{x}^* \cdot 1_{T_{j,h}} \right| \leq 2 \cdot \delta^{20} \cdot \text{OPT} \right)
\]

\[
\geq 1 - \sum_{j \in [k]} \sum_{h=2}^{2 \cdot \delta^{-1}} \Pr \left( \left| \mathbb{E} \left[ \tilde{x}^* \cdot 1_{T_{j,h}} \mid \mathcal{F}_{j-1} \right] - \tilde{x}^* \cdot 1_{T_{j,h}} \right| > 2 \cdot \delta^{20} \cdot \text{OPT} \right)
\]

\[
\geq 1 - k \cdot 2 \cdot \delta^{-1} \cdot 2 \cdot \exp \left( - \delta^{50} \cdot \text{OPT} \right)
\]

\[
\geq 1 - \delta^{-10} \cdot \exp \left( - \delta^{50} \cdot \text{OPT} \right),
\]

where the second inequality is by Lemma 3.16 and the last inequality uses \( k \leq \delta^{-2} \).
Assume the event in (10) occurs. Let \( j \in [k] \) and \( 2 \leq h \leq 2\delta^{-1} \). For any \( C \in \mathcal{C}_h \) it holds that
\[
\Pr(C \in U_{j,h} \mid \mathcal{F}_{j-1}) = \mathbb{I}_{C \in U_{j-1,h}} \cdot \Pr\left( \forall \ell \in [\rho_j] : C_\ell \cap C \cap L = \emptyset \mid \mathcal{F}_{j-1} \right)
= \mathbb{I}_{C \in U_{j-1,h}} \cdot \left( 1 - \frac{\sum_{C' \in \mathcal{C}} \mathbb{I}_{C' \cap C \cap L \neq \emptyset} x_{C'} \cdot [z_j - \ln(1-\delta)]}{z_j} \right)
\geq \mathbb{I}_{C \in U_{j-1,h}} \cdot \left( 1 - \frac{h}{z_j} \right)^{z_j - \ln(1-\delta)}
\geq \mathbb{I}_{\text{OPT}_f(\mathbb{1}_{S_{j-1}}) > \mu(\delta)} \cdot \mathbb{I}_{C \in U_{j-1,h}} \cdot \left( (1 - \delta)^h - \delta^{20} \right).
\tag{22}
\]
The first inequality holds since
\[
\sum_{C' \in \mathcal{C}} \mathbb{I}_{C' \cap C \cap L \neq \emptyset} \cdot \sum_{i \in C \cap L} C'_i = \sum_{i \in C \cap L} \sum_{C' \in \mathcal{C}} \mathbb{I}_{C' \cap C} \cdot C(i) \leq h
\]
for every \( C' \in \mathcal{C} \). The last inequality in (22) holds since \( z_j \geq \text{OPT}_f(\mathbb{1}_{S_{j-1}}) \) and by the definition of \( \mu \).

We therefore have
\[
\mathbb{1}_{U_{j,h}} \cdot \bar{x}^* = \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* - \mathbb{1}_{T_{j,h}} \cdot \bar{x}^*
\geq \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* - \mathbb{E} \left[ \mathbb{1}_{T_{j,h}} \cdot \bar{x}^* \mid \mathcal{F}_{j-1} \right] - \delta^{20} \cdot \text{OPT}
= \mathbb{E} \left[ \mathbb{1}_{U_{j,h}} \cdot \bar{x}^* \mid \mathcal{F}_{j-1} \right] - \delta^{20} \cdot \text{OPT}
\geq \text{OPT}_f(\mathbb{1}_{S_{j-1}}) > \mu(\delta) \cdot \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* \left( (1 - \delta)^h - \delta^{20} \right) - \delta^{20} \cdot \text{OPT}
\geq \text{OPT}_f(\mathbb{1}_{S_{j-1}}) > \mu(\delta) \cdot \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* \cdot (1 - \delta)^h - \delta^{19} \cdot \text{OPT}.
\tag{23}
\]

Claim 3.17. For any \( 2 \leq h \leq 2 \cdot \delta^{-1} \) and \( j \in \{0, 1, \ldots, k\} \) it holds that
\[
\bar{x}^* \cdot \mathbb{1}_{U_{j,h}} \geq (1 - \delta)^{hj} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - j \cdot \delta^{19} \cdot \text{OPT} \text{ or } \text{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta).
\]

Proof. Fix \( 2 \leq h \leq 2 \cdot \delta^{-1} \). We show the claim by induction over \( j \).

Base case: For \( j = 0 \) it clearly holds that \( \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} \geq (1 - \delta)^{h \cdot 0} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - 0 \cdot \delta^{19} \cdot \text{OPT} \).

Induction step: Assume the induction hypothesis holds for \( j - 1 \). If \( \text{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta) \) then the statement holds for \( j \). Otherwise, \( \text{OPT}_f(\mathbb{1}_{S_j}) > \mu(\delta) \) and therefore \( \text{OPT}_f(\mathbb{1}_{S_{j-1}}) \geq \text{OPT}_f(\mathbb{1}_{S_j}) > \mu(\delta) \). By the induction hypothesis, we have
\[
\bar{x}^* \cdot \mathbb{1}_{U_{j-1,h}} \geq (1 - \delta)^{h(j-1)} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - (j - 1) \cdot \delta^{19} \cdot \text{OPT}.
\tag{24}
\]

Therefore,
\[
\bar{x}^* \cdot \mathbb{1}_{U_{j,h}} \geq \text{OPT}_f(\mathbb{1}_{S_{j-1}}) > \mu(\delta) \cdot \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* \cdot (1 - \delta)^h - \delta^{19} \cdot \text{OPT}
= \mathbb{1}_{U_{j-1,h}} \cdot \bar{x}^* \cdot (1 - \delta)^h - \delta^{19} \cdot \text{OPT}
\geq (1 - \delta)^h \left( (1 - \delta)^{h(j-1)} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - (j - 1) \cdot \delta^{19} \cdot \text{OPT} \right) - \delta^{19} \cdot \text{OPT}
\geq (1 - \delta)^{hj} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - j \cdot \delta^{19} \cdot \text{OPT}.
\]

The first inequality is by (23), and the second inequality is by (24).

By Claim 3.17 for any \( j \in [k] \) and \( 2 \leq h \leq 2 \cdot \delta^{-2} \) either \( \text{OPT}_f(\mathbb{1}_{S_j}) \leq \mu(\delta) \) or
\[
\bar{x}^* \cdot \mathbb{1}_{U_{j,h}} \geq (1 - \delta)^{hj} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - j \cdot \delta^{19} \cdot \text{OPT} \geq (1 - \delta)^{hj} \cdot \bar{x}^* \cdot \mathbb{1}_{U_{0,h}} - \delta^{10} \cdot \text{OPT},
\]
as required (the last inequality uses \( j \leq k \leq \delta^{-2} \)). Since we assumed (8) occurs, this property holds with probability at least \( 1 - \delta^{-10} \cdot \exp(-\delta^{50} \cdot \text{OPT}) \) by Lemma 3.5.
We now proceed to the proof of Claim 3.8. We use the same notation as in the proof of Lemma 3.7 where the claim is stated.

**Proof of Claim 3.8**. As in the proof of Lemma 3.16 let $V \subseteq [0,1]^C$ be all the values $\bar{x}^i$ can take (formally, $V = \{\bar{x}^i(\omega) \mid \omega \in \Omega\}$). It follows that $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \leq 1$ for every $i \in I$ and $\bar{x} \in V$. Also, let $A \subseteq \mathbb{R}_{\geq 0}$ be the set of all values the vectors in $S_j$ can take (formally, $A = \{\bar{u} \mid \exists \omega \in \Omega : \bar{u} \in S_j(\omega)\}$) as $\Omega$ is finite, it follows that $V$ and $A$ are finite.

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

$$f_{U,S,\bar{x},\rho,\bar{u}}(C_1, \ldots, C_{OPT}) = \begin{cases} \frac{1}{\text{tol}(\bar{u})} \cdot \sum_{C \in U} \bar{x}_C \cdot 1_{\chi((U_{i \in [\rho]} C_i) \cap L \neq \emptyset)} \cdot \sum_{i \in C \setminus S} 1_{i \notin \bigcup_{j \in [\rho]} C_j} \cdot \bar{u}_i & \text{if } \text{tol}(\bar{u}) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

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

Let $f_{U,S,\bar{x},\rho,\bar{u}} \in D$, $(C_1, \ldots, C_{OPT}) \in c^{OPT}$ and $r \in [OPT]$ such that $C_{r} = C'_r$ for every $\ell \in [OPT] \setminus \{r\}$. If $\text{tol}(\bar{u}) = 0$ or $r > \rho$ then $|f_{U,S,\bar{x},\rho,\bar{u}}(C_1, \ldots, C_{OPT}) - f_{U,S,\bar{x},\rho,\bar{u}}(C'_1, \ldots, C'_{OPT})| = 0$. Otherwise, let $T = \bigcup_{i \notin [\rho] \setminus \{r\}} C_t = \bigcup_{i \in [\rho] \setminus \{r\}} C'_t$. Then

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

$$= \left|\frac{1}{\text{tol}(\bar{u})} \cdot \sum_{C \in U} \bar{x}_C \cdot 1_{\chi((U_{i \in [\rho]} C_i) \cap L \neq \emptyset)} \cdot \sum_{i \in C \setminus S} 1_{i \notin \bigcup_{j \in [\rho]} C_j} \cdot \bar{u}_i \right|$$

$$\leq \frac{1}{\text{tol}(\bar{u})} \cdot \sum_{C \in U} \sum_{i \in C \setminus S} \bar{x}_C \cdot \bar{u}_i \cdot \left(1_{\chi((U_{i \in [\rho]} C_i) \cap L \neq \emptyset)} - 1_{i \notin \bigcup_{j \in [\rho]} C_j} \right)$$

$$\leq \frac{1}{\text{tol}(\bar{u})} \cdot \sum_{i \in C \setminus C_{r}} \bar{x}_C \cdot \bar{u}_i \cdot \left(1_{\chi((C'_r) \cap L \neq \emptyset)} + 1_{i \in C_r \cap C'} \right)$$

$$\leq \frac{1}{\text{tol}(\bar{u})} \cdot 1_{\chi((C'_r) \cap L \neq \emptyset)} \cdot \bar{x}_C \cdot \sum_{i \in C_r \cap C'} \bar{u}_i + \frac{1}{\text{tol}(\bar{u})} \cdot \sum_{i \in C \setminus C_{r}} \bar{u}_i \cdot \sum_{C \in U} \bar{x}_C \cdot C(i)$$

$$\leq \frac{1}{\text{tol}(\bar{u})} \cdot \text{tol}(\bar{u}) \cdot 4 \cdot \delta^{-1} + \frac{1}{\text{tol}(\bar{u})} \cdot \sum_{i \in C \setminus C_{r}} \bar{u}_i$$

$$\leq 4 \cdot \delta^{-1} + \frac{1}{\text{tol}(\bar{u})} \cdot 2 \cdot \text{tol}(\bar{u})$$

$$\leq \delta^{-2},$$

where the fourth inequality uses

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

We conclude that all the functions in $D$ are of $\delta^{-2}$-bounded difference.

Recall $S_j$ is a $(\delta, \varphi(\delta))$-linear structure of $\bar{\lambda}^j$. Since $\bar{\lambda}^j$ is $F_j$-measurable, it follows that $S_j$ is also $F_j$-measurable. As in the proof of Lemma 3.4, we denote $S_j = \{\bar{u}^1, \ldots, \bar{u}^{|\varphi(\delta)|}\}$ where $\bar{u}^s$ is an $F_j$-measurable random vector for every $1 \leq s \leq |\varphi(\delta)|$ (in case $|S_j| < |\varphi(\delta)|$ the same vector may appear several times in $\bar{u}^1, \ldots, \bar{u}^{|\varphi(\delta)|}$).

For every $s \in [\varphi(\delta)]$ define a random function $g^s = f_{U_{j-1}, S_{j-1}, \bar{x}^s, \rho_j, \bar{u}^s}$. Since $U_{j-1}$, $S_{j-1}$, $\bar{x}^s$, $\rho_j$ and $\bar{u}^s$ are all $F_j$-measurable, it follows that $g$ is $F_j$-measurable as well. Furthermore,

$$\text{tol}(\bar{u}^s) \cdot g^s(C_1, \ldots, C_{OPT}) = \sum_{C \in U_{j-1}} \bar{x}_C \cdot 1_{\chi((U_{i \in [\rho]} C_i) \cap L \neq \emptyset)} \cdot \sum_{i \in C \setminus S_{j-1}} 1_{i \notin \bigcup_{j \in [\rho]} C_j} \cdot \bar{u}_i$$

$$= \sum_{i \in S_j} \bar{u}_i \cdot \sum_{C \in \mathcal{F}_j} \bar{x}_C \cdot C(i) = \sum_{i \in S_j} \bar{u}_i \cdot \bar{d}_i = (\bar{u}^s \land \bar{d}) \cdot \bar{u}^s,$$

where $\ell$ is the index such that $d = d_{\ell}$. We denote $\bar{d} = (\bar{u}^s \land \bar{d}) \cdot \bar{u}^s$. We now proceed to the proof of Claim 3.8. We use the same notation as in the proof of Lemma 3.7 where the claim is stated.
where the second equality follows from the definition of $\bar{d}^j$. Thus, for any $s \in [\varphi(\delta)]$ it holds that

$$\Pr \left( (1_{S_j} \land \bar{d}^j) \cdot \bar{u}^s > \mathbb{E} \left[ \bar{u}^s \cdot (\bar{d}^j \land 1_{S_j}) \mid \mathcal{F}_{j-1} \right] + \frac{\text{OPT}}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}) \right) = \Pr \left( g^s(C_1^j, \ldots, C_{OPT}^j) > \mathbb{E} \left[ g^s(C_1^j, \ldots, C_{OPT}^j) \mid \mathcal{F}_{j-1} \right] + \frac{\text{OPT}}{\varphi^{11}(\delta)} \right) \leq \exp \left( -\frac{2 \cdot \left( \frac{\text{OPT}}{\varphi^{11}(\delta)} \right)^2}{\delta^{-4} \cdot \text{OPT}} \right) \leq \exp \left( -\frac{\text{OPT}}{\varphi^{25}(\delta)} \right),$$

where the last inequality is by Lemma 3.13.

Thus, using the union bound we have that

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

It remains to prove Lemma 3.3 and Claim 3.11. We use $G = (L, E)$ to denote the $\delta$-matching graph of $(I, v)$, and $P_M(G)$ to denote the matching polytope of $G$. Both proofs rely on concentration bounds of SampleMatching given below.

**Lemma 3.18 (H).** Let $\beta \in P_M(G)$ and $\gamma > 0$. Also, denote $M = \text{SampleMatching}(\beta, \gamma)$. Then $M$ is a matching and for any $\bar{a} \in [0, 1]^E$ the following holds:

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

**Proof of Lemma 3.3** As $M = \text{SampleMatching}(\mathcal{E}(\bar{x}^0), \delta^4)$, it follows that

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

If $\mathbb{1}_{C} \cdot \bar{x}^0 = 0$ then $|M| = 0$ and the statement of the lemma holds. Otherwise, by Lemma 3.18

$$\Pr \left( |M| > \mathbb{1}_{C} \cdot \bar{x}^0 + \delta^2 \cdot \text{OPT} \right) = \Pr \left( |M| > \mathbb{1}_{C} \cdot \bar{x}^0 \cdot \left( 1 + \frac{\delta^2 \cdot \text{OPT}}{\mathbb{1}_{C} \cdot \bar{x}^0} \right) \right) \leq \exp \left( -\frac{1}{20} \cdot \delta^4 \cdot (\mathbb{1}_{C} \cdot \bar{x}^0) \cdot \left( \frac{\delta^2 \cdot \text{OPT}}{\mathbb{1}_{C} \cdot \bar{x}^0} \right)^2 \right) \leq \exp \left( -\delta^{10} \cdot \text{OPT} \right),$$

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

$$\Pr \left( |M| \leq \mathbb{1}_{C} \cdot \bar{x}^0 + \delta^2 \cdot \text{OPT} \right) \geq 1 - \exp \left( -\delta^{10} \cdot \text{OPT} \right).$$

□
Proof of Claim 3.11. We use the same notation as in the proof of Lemma 3.10, where the claim is stated. If tol(\(\bar{u}\)) = 0 the claim trivially holds. Thus, we may assume that tol(\(\bar{u}\)) ≠ 0.

Observe that
\[
\bar{w} \cdot \bar{u} - \bar{d} \cdot \bar{u} = \sum_{i \in I} (\bar{w}_i - \bar{d}_i) \bar{u}_i = \sum_{i \in L} (1 - \delta^4) \cdot \bar{y}_i^M \bar{u}_i = \sum_{i \in L} 1_{i \notin S_0} \cdot \bar{u}_i - E \left[ \sum_{i \in L} 1_{i \notin S_0} \cdot \bar{u}_i \right],
\]
where the second equality is by (16) and (17), and the last equality is by Lemma 3.9. Furthermore,
\[
\sum_{i \in L} 1_{i \notin S_0} \cdot \bar{u}_i = \sum_{\{i_1, i_2\} \in M} (\bar{u}_{i_1} + \bar{u}_{i_2})
\]
Thus,
\[
\Pr \left( \bar{d} \cdot \bar{u} > \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}) \right) = \Pr \left( \sum_{i \in L} 1_{i \notin S_0} \cdot \bar{u}_i < E \left[ \sum_{i \in L} 1_{i \notin S_0} \cdot \bar{u}_i \right] - \frac{\text{OPT}}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}) \right) = \Pr \left( \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} < E \left[ \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} \right] - \frac{\text{OPT}}{\varphi^{11}(\delta)} \right) \leq \exp \left( -\frac{1}{20} \cdot \delta^4 \cdot \frac{\text{OPT}}{\varphi^{25}(\delta)} \right).
\]
The first inequality is by Lemma 3.18 observe that \(M \subseteq E \subseteq C\), therefore \(\frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} \leq 1\) for any \(\{i_1, i_2\} \in E\). The last inequality uses
\[
E \left[ \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} \right] \leq \frac{|L|}{2} \leq \delta^{-1} \cdot \text{OPT}.
\]

It is implicitly assumed in (25) that \(E \left[ \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} \right] \neq 0\). In case \(E \left[ \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} \right] = 0\), we have \(\sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} = 0\), and
\[
\Pr \left( \bar{d} \cdot \bar{u} > \bar{w} \cdot \bar{u} + \frac{\text{OPT}}{\varphi^{11}(\delta)} \cdot \text{tol}(\bar{u}) \right) = \Pr \left( \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} < E \left[ \sum_{\{i_1, i_2\} \in M} \frac{\bar{u}_{i_1} + \bar{u}_{i_2}}{\text{tol}(\bar{u})} \right] - \frac{\text{OPT}}{\varphi^{11}(\delta)} \right) = \Pr \left( 0 < -\frac{\text{OPT}}{\varphi^{25}(\delta)} \right) = 0 \leq \exp \left( -\frac{\text{OPT}}{\varphi^{25}(\delta)} \right).
\]

4 Proof of the Structural Lemma

In this section we give the proof of Lemma 2.2 Let \(\delta \in (0, 0.1)\) such that \(\delta^{-1} \in \mathbb{N}\), and let \((I, v)\) be a \(\delta\)-2VBP instance. As in Section 3 we use \(\text{OPT} = \text{OPT}(I, v)\).

We first need to construct the set \(S^* \subseteq \mathbb{R}_{\geq 0}^I\). 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 \(S^*\) are also used in the construction of the linear structure \(S\).

Let \(\succeq^*\) be an arbitrary total order 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 any \(d \in \{1, 2\}\) and
Let $\bar{\lambda} \in [0, 1]^{C^*}$ be a small items integral vector with $\delta$-slack, and let $\bar{w} \in [0, 1]^I$ be the coverage of $\bar{\lambda}$. In Section 4.1 we construct the linear structure $S$ of $\bar{\lambda}$, and in Section 4.2 we show the structure indeed satisfies the requirements in Definition 2.1. The construction and proof of correctness rely on a technical refinement lemma whose proof is given in Section 4.3.

**4.1 Construction of $S$**

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 $C \in \text{supp}(\bar{\lambda}^d)$ it holds that $C$ has $\delta$-slack in dimension $d$. Formally, we define $\bar{\lambda}^1 \in [0, 1]^{C^*}$ by

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

Also, we define $\bar{\lambda}^2 \in [0, 1]^{C^*}$ by $\bar{\lambda}^2 = \bar{\lambda} - \bar{\lambda}^1$. Indeed, as $\bar{\lambda}$ is with $\delta$-slack, for every $d \in \{1, 2\}$ and $C \in \text{supp}(\bar{\lambda}^d)$, it holds that $C$ has $\delta$-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 $j \cdot \frac{2}{\ell}$. The slack of configurations in $\text{supp}(\bar{\lambda})$ is used to compensate for the possible volume increase. In the other dimension, $d$, we apply fractional grouping, defined as follows.

**Definition 4.1.** Let $E \neq \emptyset$ be an arbitrary 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 $\xi$-fractional grouping w.r.t $\bar{\gamma}$ 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 every $\ell \in [\tau - 1]$ it holds that $1_{G_{\ell}} \cdot \bar{\gamma} \geq \frac{\|\bar{\gamma}\|}{\xi}$.
3. For every $\ell \in [\tau]$ it holds that $1_{G_{\ell}} \cdot \bar{\gamma} \leq \frac{\|\bar{\gamma}\|}{\xi} + 1$.

The proof of the next lemma utilizes arguments from [12].

**Lemma 4.2.** 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 $\xi$-fractional grouping $G_1, \ldots, G_{\tau}$ of $E$ w.r.t $\bar{\gamma}$ and $\succeq$ for which $\tau \leq \xi$.

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

W.l.o.g assume $E = \{1, 2, \ldots, \nu\} = [\nu]$ and $a \succeq b$ if and only if $a \leq b$. Define a sequence $(q_{\ell})_{\ell=0}^{\infty}$ by $q_0 = 0$, and $q_{\ell} = \min\left\{e \in E \mid \sum_{f=q_{\ell-1}+1}^{q_{\ell}} \bar{\gamma}_f > \frac{\|\bar{\gamma}\|}{\xi} \right\}$ or $\nu$. Also, define $\tau = \min\{\ell \in \mathbb{N} \mid q_{\ell} = \nu\}$.

Since $\|\bar{\gamma}\| > 0$, it follows that $(q_{\ell})_{\ell=0}^{\infty}$ is monotonically increasing.

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

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

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

---

\[\text{We refer the reader to Appendix B.2 of [3] for a formal definition of total order.}\]
Thus, we showed that $G_1, \ldots, G_\tau$ is a $\xi$-fractional grouping of $E$ w.r.t $\gamma$ and $\succeq$. It also holds that
\[
\|\gamma\| = \sum_{e \in E} \bar{\gamma}_e = \tau \sum_{\ell=1}^\tau \sum_{e \in G_{\ell}} \bar{\gamma}_e \geq \tau \sum_{\ell=1}^{\tau-1} \sum_{j \in \tau e} \bar{\gamma}_e > \tau \sum_{\ell=1}^{\tau-1} \|\bar{\gamma}\|_{\xi} = (\tau - 1) \|\bar{\gamma}\|_{\xi}.
\]
Hence, $\tau - 1 < \xi$, and as both $\tau$ and $\xi$ are integral it follows that $\tau \leq \xi$. This completes the proof. \qed

For any $d \in \{1, 2\}$ and $j \in [2h]$ define a vector $\bar{\gamma}^{d,j} \in [0, 1]^I_{d,j}$ by $\bar{\gamma}^{d,j} = \bar{w}_d$ for $i \in I_{d,j}$. By Lemma 4.2 for any $d \in \{1, 2\}$ and $j \in [2h]$ such that $I_{d,j} \neq \emptyset$ there is an $h$-fractional grouping $(G^\delta_{\ell})_{\ell=1}^{\tau_{d,j}}$ of $I_{d,j}$ w.r.t $\delta_{d,j}$ and the total order $\succeq_{d,j}$ with $\tau_{d,j} \leq h$. 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 $G_1, G_2 \subseteq [2h] \times [h]$ and thus $|G_1|, |G_2| \leq 2h^4$.

Our objective is to add to the structure $S$ vectors $\hat{a}$ to ensure that if $\vec{z} \in [0, 1]^I$ satisfies (8) then we can decompose $\vec{z} \wedge 1_L$ to $\vec{z}^1, \vec{z}^2 \in [0, 1]^I$ such that $\vec{z} = \vec{z}^1 + \vec{z}^2$ and $\vec{z}^1 \cdot G^{d,j}_{\ell} \preceq \beta \cdot \bar{w}^d \cdot \bar{G}^{d,j}_\ell$ for any $d \in \{1, 2\}$ and $(j, \ell) \in \mathcal{G}_d$. This can intuitively be interpreted as a decrease in demand for items in $G^{d,j}_\ell$ by a factor of $\beta$.

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 $S$ vectors which represent the intersection of each pair of such groups, and therefore impose a decrease in demand by a factor of $\beta$ for each intersection.

Formally, our linear structure will contain the set $S_{\text{large}}$ defined below.

$$
S_{\text{large}} = \left\{ \mathbf{1}_{G_{\ell_1}^{1,j_1}} \wedge \mathbf{1}_{G_{\ell_2}^{2,j_2}} \mid (j_1, \ell_1) \in G_1, (j_2, \ell_2) \in G_2 \right\}.
$$

(26)

In Section 4.2 we show that if $S_{\text{large}} \subseteq S$ and $\vec{z}$ satisfies (9) then we can find the decomposition $\vec{z}^1$ and $\vec{z}^2$ as mentioned above. Furthermore, to show the correctness of the structure we (implicitly) use a shifting argument (see, e.g., [21]) 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 $T^d(C)$, is the vector $\bar{t} \in \mathbb{N}^{G_d}$ 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 $\bar{t}(j, \ell) \leq 2h^4$. Let $T_d = \{T^d(C) \mid C \in \mathcal{C}\}$ be the set of all possible $d$-types. It follows that $T_d \subseteq \{0, 1, \ldots, 2 \cdot \delta^{-1}\}^{G_d}$, and therefore $|T_d| \leq (1 + 2 \cdot \delta^{-1})^{2 \delta^{-4}} \leq \exp(\delta^{-6})$.

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

$$
\bar{a}_{d,\bar{t}}(i) = \sum_{C \in \mathcal{C} \text{ s.t. } T^d(C) = \bar{t}} \bar{\lambda}_C \cdot C(i),
$$

(27)

for $i \in I \setminus L$ and $\bar{a}_{d,\bar{t}} = 0$ 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 $\bar{a}^d$.

For $d \in \{1, 2\}$ define $\bar{\bar{a}}^d \in [0, 1]^I$ by $\bar{\bar{a}}^d(i) = a_d(i)$ 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 4.3** (Small Items Refinement). Let $\bar{a} \in [0, 1]^I$ such that $\text{supp}(\bar{a}) \subseteq I \setminus L$, $d \in \{1, 2\}$ and $q \in \mathbb{N}$, $q \geq 4$. Given $\beta \in \left[\frac{1}{q}, 1\right]$, there are subsets $H_1, \ldots, H_q \subseteq I \setminus L$ such that for any $Q \subseteq I \setminus L$ satisfying

$$
\forall 1 \leq j \leq q : \quad \|\mathbf{1}_{Q \cap H_j} \bullet \bar{a} \bullet \bar{\bar{a}}^d\| \leq \beta \|\mathbf{1}_{H_j} \bullet \bar{a} \bullet \bar{\bar{a}}^d\| + \frac{\text{OPT}}{q^5} \max \{v_d(C \cap H_j) \mid C \in \mathcal{C}\},
$$

(28)

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

1. $\|\mathbf{1}_X \bullet \bar{a} \bullet (\bar{\bar{a}}^1 + \bar{\bar{a}}^2)\| \leq \frac{16}{q} \cdot \text{OPT} + 2q\delta$.  

(27)
2. \[ \|\hat{a} \cdot \bar{a} \cdot \hat{v}^d\| \leq \beta \cdot \bar{a} \cdot \hat{v}^d. \]

We refer to \(H_1, \ldots, H_q\) as the refinement of \(\bar{a}\) and \(q\) in dimension \(d\).

Indeed, the condition in (28) is essentially a variant of (5). Lemma 4.3 plays a central role in showing the correctness of the structure \(S\). (see the proof of Lemma 4.2). We defer the proof of Lemma 4.3 to Section 4.3.

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

\[
S_{\text{small}} = \left\{ \hat{1}_{H^d_t, \ldots, H^{d,d'}_q} \cdot \bar{a}^{d,j} \cdot \hat{v}^d \mid d, d' \in \{1, 2\}, \bar{t} \in \mathcal{T}_d, j \in [q] \right\}.
\]

Finally, the structure is \(S = S_{\text{large}} \cup S_{\text{small}}\).

### 4.2 Correctness

We first observe that

\[
|S| = |S_{\text{large}}| + |S_{\text{small}}| \leq |G_1| + |G_2| + 2 \cdot q \cdot (|\mathcal{T}_1| + |\mathcal{T}_2|) \leq \exp(\delta^{-20}) = \varphi(\delta).
\]

Let \(\bar{u} \in S\) such that \(\text{supp}(\bar{u}) \cap L \neq \emptyset\), then \(\bar{u} \in S_{\text{large}}\). Therefore, by (20) there is \((j_1, \ell_1) \in G_1\) and \((j_2, \ell_2) \in G_2\) such that \(\bar{u} = \hat{1}_{G_{j_1}^{1,j_1}} \land \hat{1}_{G_{\ell_2}^{2,\ell_2}}\). By Definition 4.1, for \(d \in \{1, 2\}\) there are \(q_d^1, q_d^2 \in I_{d,j_d}\) such that \(G^d_{d,j_d} = \left\{ i \in I_{d,j_d} \mid q_d^1 \leq i \leq q_d^2 \right\}\); thus, \(\hat{1}_{G^d_{j_d,k_d}} \in S^*_d\). It follows that \(\bar{u} = \hat{1}_{G_{j_1}^{1,j_1}} \land \hat{1}_{G_{\ell_2}^{2,\ell_2}} \in S^*\).

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

\[
\forall \bar{u} \in S : \quad \bar{z} \cdot \bar{u} \leq \beta \cdot \bar{w} \cdot \bar{u} + \frac{1}{\varphi^{10}(\delta)} \cdot \text{OPT} \cdot \text{tol}(\bar{u}) \tag{29}
\]

To verify that \(S\) is a \((\delta, \varphi(\delta))\) linear structure, it remains to show that \(\text{OPT}_{f}(\bar{z}) \leq \beta(1 + 10\delta) \cdot \|\bar{\lambda}\| + \varphi(\delta) + \delta^{10} \cdot \text{OPT}(I, v)\).

We first generate two vectors \(\bar{z}^1\) and \(\bar{z}^2\) such that \(\bar{z} \land 1_L \land \bar{z}^d_{\ell} \cdot \bar{1}_{G^d_{j_d,k_d}} \leq \beta \bar{w}^d_{\ell} \cdot \bar{1}_{G^d_{j_d,k_d}}\) for every \(d \in \{1, 2\}\) and \((j, \ell) \in G_d\). Each item \(i \in L\) belongs to groups \(G_{j_1}^{1,j_1}\) and \(G_{\ell_2}^{2,\ell_2}\). 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 \(G_{j_1}^{1,j_1} \land G_{\ell_2}^{2,\ell_2}\). Specifically, for \(d \in \{1, 2\}\), define \(\bar{z}^d \in [0, 1]^I\) by

\[
\forall (j_1, \ell_1) \in G_1, (j_2, \ell_2) \in G_2, i \in G_{j_1}^{1,j_1} \land G_{\ell_2}^{2,\ell_2} \land \text{supp}(\bar{z}) : \quad \bar{z}_i^d = \frac{1}{(1_{G_{j_1}^{1,j_1}} \land 1_{G_{\ell_2}^{2,\ell_2}}) \cdot \bar{w}} \cdot \frac{1_{G_{j_1}^{1,j_1}} \land \bar{z}^d_{\ell} \cdot 1_{G_{\ell_2}^{2,\ell_2}}}{1_{G_{j_1}^{1,j_1}} \land 1_{G_{\ell_2}^{2,\ell_2}}} \tag{30}
\]

and \(\bar{z}_i = 0\) for any other \(i \in I\). Observe that since \(\text{supp}(\bar{z}) \subseteq \text{supp}(\bar{w})\) we never get in (30) a division by zero. Since for every \(i \in L\) there is a unique \((j_1, \ell_1) \in G_1\) and a unique \((j_2, \ell_2) \in G_2\) such that \(i \in G_{j_1}^{1,j_1} \land G_{\ell_2}^{2,\ell_2}\), it follows that \(\bar{z} \land 1_L = \bar{z}^1 + \bar{z}^2\). For every \(d \in \{1, 2\}\) and \((j, \ell) \in G_d\) it holds that

\[
\bar{z}^d \cdot 1_{G_{j_d}^{d,j_d}} = \sum_{(j', \ell') \in G_d} \sum_{i \in G_{j_d}^{d,j_d} \land 1_{G_{j_d}^{d,j_d}} \land \text{supp}(\bar{z})} \bar{z}_i \cdot \frac{1_{G_{j_d}^{d,j_d}} \land 1_{G_{j_d}^{d,j_d}} \cdot \bar{w}}{1_{G_{j_d}^{d,j_d}} \land 1_{G_{j_d}^{d,j_d}} \cdot \bar{w}} = \sum_{(j', \ell') \in G_d, \text{s.t.} \left(1_{G_{j_d}^{d,j_d}} \land 1_{G_{j_d}^{d,j_d}} \cdot \bar{w} \neq 0\right)} \left(1_{G_{j_d}^{d,j_d}} \land 1_{G_{j_d}^{d,j_d}} \cdot \bar{w}ight) \cdot \bar{z}_i \cdot \frac{1_{G_{j_d}^{d,j_d}} \land 1_{G_{j_d}^{d,j_d}} \cdot \bar{w}}{1_{G_{j_d}^{d,j_d}} \land 1_{G_{j_d}^{d,j_d}} \cdot \bar{w}} \tag{31}
\]
Since \(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}} \in S\), by (29)

\[
\left(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}}\right) \cdot \bar{\varepsilon} \leq \beta \left(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}}\right) \cdot \bar{\omega} + \frac{\text{OPT}}{\phi^{10}(\delta)} \cdot \text{tol} \left(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}}\right)
\]

\[
\leq \beta \left(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}}\right) \cdot \bar{\omega} + \frac{2 \cdot \delta^{-1} \cdot \text{OPT}}{\phi^{10}(\delta)}.
\]

The second inequality holds since there are at most \(2\delta^{-1}\) large items in a configuration. Plugging (32) into (31) we have

\[
\bar{\varepsilon}^d \cdot 1_{G^d_{i,j}}
\]

\[
\leq \sum_{(j',\ell') \in \mathcal{G}_d \text{ s.t. } (C^d_{i,j} \wedge 1_{G^d_{i,j'}}) \cdot \bar{w} \neq 0} \beta \left(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}}\right) \cdot \bar{\omega}^d + \frac{\delta^{-6}}{\phi^{10}(\delta)} \cdot \text{OPT},
\]

where the second inequality holds since \(|G_d| \leq 2 \cdot \delta^{-4}\).

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

\[
(\bar{z}^d - \bar{r}^d) \cdot 1_{C^d_{i,j}} \leq \max \left\{ \beta \cdot 1_{C^d_{i,j}} \cdot \bar{w}^d - 2, 0 \right\},
\]

for every \(i \in I\) it holds that \(r^d_i \leq z^d_i\), and \(\|\bar{r}^d\| \leq \left(2 + \frac{\delta^{-6}}{\phi^{10}(\delta)} \cdot \text{OPT} \right) \cdot |G_d| \leq \delta^{-5} + \frac{\delta^{-11}}{\phi^{10}(\delta)} \cdot \text{OPT}\). Hence, \(\text{OPT}_f(\bar{r}^d) \leq \|\bar{r}^d\| \leq \delta^{-5} + \frac{\delta^{-11}}{\phi^{10}(\delta)} \cdot \text{OPT}\), as \(\sum_{i \in I} r^d_i \cdot 1_{\{i\}}\) is a solution for LP(\(\bar{r}^d\)).

For any \(d \in \{1, 2\}\), let \(F_d = \bigcup_{j \in [2h]} \text{ s.t. } (j, 1) \in \mathcal{G}_d\) be the set of all items that belong to a first group in one of the fractional groupings \(G^d_{i,j}, \ldots, G^d_{i,j'}\). By (31),

\[
(\bar{z}^d - \bar{r}^d) \cdot 1_{F_d} \leq \sum_{j \in [2h]} \left(1_{C^d_{i,j}} \wedge 1_{G^d_{i,j'}}\right) \cdot \bar{w}^d \leq \sum_{j \in [2h]} \max \left\{ \beta \cdot 1_{C^d_{i,j}} \cdot \bar{w}^d - 2, 0 \right\}
\]

\[
\leq \beta \sum_{j \in [2h]} \max \left\{ \bar{w}^d \cdot 1_{L_j} \right\} = \frac{\bar{\omega}^d \cdot \mathbf{1}_L}{h} \leq 2 \cdot \delta \cdot \beta \cdot \|\bar{\lambda}^d\|
\]

where the third inequality is by Definition (31) and the last inequality follows from \(h = \delta^{-2}\) and

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

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

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

Then,

\[
\text{OPT}_f(\bar{z}) \leq \sum_{d \in \{1, 2\}} \left(\text{OPT}_f(\bar{r}^d) + \text{OPT}_f((\bar{z}^d - \bar{r}^d) \wedge 1_{F_d})\right) + \text{OPT}_f(\bar{y})
\]

\[
\leq \text{OPT}_f(\bar{y}) + 2\beta\delta\|\bar{\lambda}\| + 2\delta^{-5} + \frac{2 \cdot \delta^{-11}}{\phi^{10}(\delta)} \cdot \text{OPT}.
\]

(36)
We proceed to derive an upper bound on \( \text{OPT}_f(\tilde y) \), which in turn implies an upper bound on \( \text{OPT}_f(\tilde z) \).

Given \( d \in \{1, 2\} \) we define the \( d \)-size of \((j, \ell) \in G_d \), denoted \( s^d(j, \ell) \in [0, 1]^2 \), by \( s^d_0(j, \ell) = \frac{d^2}{2} j \) and \( s^d_1 = \min \{ v_d(i) \mid i \in G^d_{\ell,j} \} \). The value \( s^d(j, \ell) \) can be viewed as the rounded volume of items in \( G^d_{\ell,j} \).

The next lemma gives the basis for our shifting argument.

**Lemma 4.4.** Let \( d \in \{1, 2\} \), \((j, \ell) \in G_d \) and \( i \in G^d_{\ell,j} \). If \( \ell \neq 1 \) then \( v(i) \leq s^d(j, \ell - 1) \).

**Proof.** As \( i \in G^d_{\ell,j} \subseteq I_{d,j} \), it follows that \( v_d(i) \leq \frac{d^2}{2} j = s^d_0(j, \ell - 1) \). Furthermore, \( v_d(i') \geq v_d(i) \) for every \( i' \in G^d_{\ell-1,j} \) as \( (G^d_{\ell,j})'_{\ell=1} \) is an \( h \)-fractional grouping w.r.t to the relation \( \preceq \). Hence, \( v_d(i) \leq \min \{ v_d(i') \mid i' \in G^d_{\ell-1,j} \} = s^d(j, \ell - 1) \).

We extend the definition of size to \( d \)-types by \( s^d(\ell) = \sum_{(j, \ell) \in G_d} \ell(j, \ell) \cdot s^d(j, \ell) \) for any \( d \in \{1, 2\} \) and \( \ell \in T_d \).

**Lemma 4.5.** Let \( d \in \{1, 2\} \) and \( C \in C^* \) with \( \hat{\lambda}_d^C > 0 \). Then \( \sum_{i \in I \setminus L} v(i) \cdot C(i) \leq 1 - s^d(T^d(C)) \).

**Proof.** For any \( i \in L \) such that \( C(i) > 0 \) there is a unique \((j, \ell) \in G_d \) for which \( i \in G^d_{\ell,j} \). Thus,

\[
\sum_{i \in I \setminus L} v(i) \cdot C(i) = \sum_{i \in I} v(i) \cdot C(i) - \sum_{i \in L} v(i) \cdot C(i) = v(C) - \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} v(i) \cdot C(i) \quad (37)
\]

Therefore, we have

\[
\sum_{i \in I \setminus L} v_d(i) \cdot C(i) = v_d(C) - \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} v_d(i) \cdot C(i)
\]

\[
\leq 1 - \delta - \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} \left( s^d_0(j, \ell) - \frac{\delta^2}{2} \right) \cdot C(i)
\]

\[
= 1 - \delta - \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} s^d_0(j, \ell) \cdot C(i) + \frac{\delta^2}{2} \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} C(i) \quad (38)
\]

\[
= 1 - \delta - \sum_{(j, \ell) \in G_d} T^d_{(j, \ell)}(C) \cdot s^d_0(j, \ell) + \frac{\delta^2}{2} \sum_{i \in L} C(i)
\]

\[
\leq 1 - s^d(T^d(C)).
\]

The first equality is by (37). The first inequality holds as \( C \) has \( \delta \)-slack in dimension \( d \) since \( \hat{\lambda}_d^C > 0 \), and since \( v_d(i) > \frac{\delta^2}{2} (j - 1) \) for any \( i \in G^d_{\ell,j} \). The last inequality holds as there are at most \( 2\delta^{-1} \) large items in a multi-configuration. Similarly,

\[
\sum_{i \in I \setminus L} v_d(i) \cdot C(i) = v_d(C) - \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} v_d(i) \cdot C(i)
\]

\[
\leq 1 - \sum_{(j, \ell) \in G_d} \sum_{i \in G^d_{\ell,j}} s^d_1(j, \ell) \cdot C(i) \quad (39)
\]

\[
= 1 - \sum_{(j, \ell) \in G_d} T^d_{(j, \ell)}(C) \cdot s^d_1(j, \ell)
\]

\[
\leq 1 - s^d_1(T^d(C)).
\]

The first equality follows from (37) and the first inequality is by the definition of \( s^d_1(j, \ell) \). The statement of the lemma follows from (38) and (39).
For any \( d \in \{1, 2\} \) and \( \bar{t} \in T_d \), the prevalence of type \( \bar{t} \) is \( \eta_d(\bar{t}) = \sum_{C \in C^*} \tau^d(C) = \bar{\lambda}^d \). Informally, \( \eta_d(\bar{t}) \) is the number of configurations of type \( \bar{t} \) selected by \( \bar{\lambda}^d \). Also, define \( \kappa_d(\bar{t}) = [\beta \cdot \eta_d(\bar{t})] + 2 \cdot \delta^{-1} \) for any \( d \in \{1, 2\} \) and \( \bar{t} \in T_d \). 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 4.6.** There are vectors \( \bar{x}^{d, \bar{t}} \in [0, 1]^C \) for \( d \in \{1, 2\} \) and \( \bar{t} \in T_d \) such that

1. For any \( d \in \{1, 2\} \) the coverage of \( \sum_{\bar{t} \in T_d} \kappa_d(\bar{t}) \cdot \bar{x}^{d, \bar{t}} \) is \( (\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d} \).
2. For any \( d \in \{1, 2\} \) and \( \bar{t} \in T_d \) it holds that \( \|\bar{x}^{d, \bar{t}}\| = 1 \).
3. For any \( d \in \{1, 2\} \), \( \bar{t} \in T_d \) and \( C \in \text{supp}(\bar{x}^{d, \bar{t}}) \), it holds that \( v(C) \leq s^d(\bar{t}) \).

The proof of Lemma 4.6 relies on the following combinatorial claim (we omit the proof).

**Claim 4.7.** Let \( E \) be an arbitrary finite set, \( \xi \in \mathbb{Z}_+ \) 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 4.6.** Let \( d \in \{1, 2\} \) and for any \( (j, \ell) \in G_d \), define \( \rho(j, \ell) = \sum_{\bar{t} \in T_d} \bar{t}(j, \ell) \cdot \kappa_d(\bar{t}) \). Then \( \rho(j, \ell) \geq 2 \cdot \delta^{-1} \). For any \( (j, \ell) \in G_d \) and \( i \in G^{d, j}_\ell \) such that \( \ell \neq 1 \) define \( p_i = \frac{\bar{z}_{d,j}^d - \eta_d(\bar{t})}{\rho(j, \ell - 1)} \leq \frac{1}{2 \delta^{-1}} \).

For every \( (j, \ell) \in G_d \) with \( \ell \neq 1 \) it holds that

\[
\rho(j, \ell - 1) = \sum_{\bar{t} \in T_d} \bar{t}(j, \ell - 1) \cdot \kappa_d(\bar{t}) \geq \sum_{\bar{t} \in T_d} \bar{t}(j, \ell - 1) \cdot \eta_d(\bar{t}) = \beta \sum_{\bar{t} \in T_d} \bar{w}^d \cdot \mathbb{1}_{G^{d,j}_\ell} \geq \beta \frac{\bar{w}^d \cdot \mathbb{1}_{I_{d,j}}}{h} \geq \max \left\{ \beta \cdot \bar{w}^d \cdot \mathbb{1}_{G^{d,j}_\ell} - 1, 0 \right\} \geq \left( \bar{z}^d - \bar{r}^d \right) \cdot \mathbb{1}_{G^{d,j}_\ell}.
\]

The second and third inequalities hold since \( G^{d,j}_1, \ldots, G^{d,j}_n \) is an \( h \)-fractional grouping of \( I_{d,j} \). The last inequality is by (51). Therefore \( \sum_{i \in G^{d,j}_\ell} p_i \leq 1 \).

Fix \( \bar{t} \in T_d \), and for any \( (j, \ell) \in G_d \) with \( \ell \neq 1 \) let \( K(j, \ell) \subseteq G^{d,j}_\ell \) be a random set such that \( |K(j, \ell)| \leq \tilde{t}(j, \ell - 1) \) and \( \Pr(i \in K(j, \ell)) = \tilde{t}(j, \ell - 1) \cdot p_i \) for every \( i \in G^{d,j}_\ell \). The random sets \( K(j, \ell) \) exist by Claim 4.7. Furthermore, we may assume the random sets \( K(j, \ell) \) are independent.

Define \( R = \bigcup_{(j, \ell) \in G_d} K(j, \ell) \) and \( \bar{x}^{d, \bar{t}} = \Pr(R = C) \) for all \( C \in C^* \). It follows that \( \|\bar{x}^{d, \bar{t}}\| = \sum_{C \in C^*} \Pr(R = C) = 1 \). Observe that

\[
v(R) \leq \sum_{(j, \ell) \in G_d} v(K(j, \ell)) \leq \sum_{(j, \ell) \in G_d} \tilde{t}(j, \ell - 1) \cdot s^d(j, \ell - 1) \leq s^d(\bar{t}).
\]

The second inequality holds since \( |K(j, \ell)| \leq \tilde{t}(j, \ell - 1) \) and for every \( i \in K(j, \ell) \) it holds that \( v(i) \leq s^d(j, \ell - 1) \) by Lemma 4.5. 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 \( i \in \text{supp}((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}) \), there is \( (j, \ell) \in G_d \) with \( \ell \neq 1 \) such that \( i \in G^{d,j}_\ell \). Hence,

\[
\sum_{C \in C^*} \bar{x}^{d, \bar{t}} \cdot C(i) = \Pr(i \in R) = \tilde{t}(j, \ell - 1) \cdot \frac{\bar{z}_{d,j}^d - \bar{r}_{d,1}^d}{\rho(j, \ell - 1)} \geq \sum_{C \in C^*} \bar{x}^{d, \bar{t}} \cdot C(i).
\]

Let \( \bar{w}^d \) be the coverage of \( \sum_{\bar{t} \in T_d} \kappa_d(\bar{t}) \cdot \bar{x}^{d, \bar{t}} \). By construction, we have \( \bar{w}^d = 0 \) for any \( i \in I \) such that \( i \notin \text{supp}((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}) \). For any \( i \in \text{supp}((\bar{z}^d - \bar{r}^d) \wedge \mathbb{1}_{L \setminus F_d}) \), it holds that

\[
\bar{w}^d(i) = \sum_{C \in C^*} \sum_{\bar{t} \in T_d} \kappa_d(\bar{t}) \cdot \bar{x}^{d, \bar{t}} \cdot C(i) = \sum_{\bar{t} \in T_d} \kappa_d(\bar{t}) \cdot \tilde{t}(j, \ell - 1) \cdot \frac{\bar{z}_{d,j}^d - \bar{r}_{d,1}^d}{\rho(j, \ell - 1)} = \frac{\bar{z}_{d,j}^d - \bar{r}_{d,1}^d}{\rho(j, \ell - 1)} \geq \frac{\bar{z}_{d,j}^d - \bar{r}_{d,1}^d}{\rho(j, \ell - 1)}.
\]

where the second equality is by (40), 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
\begin{equation}
B = \{(d, \bar{t}, m) \mid d \in \{1, 2\}, \bar{t} \in T_d, m \in [\kappa_d(\bar{t})]\} \cup [M].
\end{equation}
We consider $B$ as a set of bins, and define a polytope
\begin{equation}
P = \left\{ \bar{\mu} \in [0, 1]^{Q \times B} \mid \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 1 - s^d(\bar{t}) & \forall d \in \{1, 2\}, \bar{t} \in T_d, m \in [\kappa_d(\bar{t})] \\
\sum_{i \in Q} \bar{\mu}_{i,m} \cdot v(i) \leq 1 & \forall m \in M
\end{array}\right\}
\end{equation}
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 (41) 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., [2]).

**Lemma 4.8.** 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 4.8 we need to show that $P$ has a vertex.

**Lemma 4.9.** $P \neq \emptyset$.

**Proof.** Ideally, we would like to define $\bar{\mu}_{i,(d,\bar{t},m)} = \frac{a_{d,\bar{t}}}{\kappa_d(\bar{t})}$ for any $i \in Q$, $d \in \{1, 2\}$, $\bar{t} \in T_d$ and $m \in [\kappa_d(\bar{t})]$. Using (29) we can show that $\sum_{i \in Q} \bar{\mu}_{i,(d,\bar{t},m)} \cdot v_d(i)$ is not significantly larger than $1 - s^d(\bar{t})$; however, we cannot show it is smaller (or equal) to $1 - s^d(\bar{t})$. Thus, the suggested vector $\bar{\mu}$ may not satisfy the properties in (41). We use Lemma 4.3 to overcome this difficulty. Specifically, we define $\bar{\mu}_{i,(d,\bar{t},m)} = \frac{a_{d,\bar{t}}}{\kappa_d(\bar{t})}$ for items $i \in Q \setminus X_1 \setminus X_2$, where the sets $X_1$ and $X_2$ are obtained via Lemma 4.3. The value of $\bar{\mu}_{i,m}$ is subsequently increased for $i \in X_1 \cup X_2$ to ensure the first constraint in (41) holds. Property 11 of Lemma 4.3 is used to show that $\sum_{i \in Q} \bar{\mu}_{i,m} \cdot v(i) \leq 1$, and property 2 of the lemma is used to show that $\sum_{i \in Q} \bar{\mu}_{i,(d,\bar{t},m)} \cdot v_d(i) \leq 1 - s^d(\bar{t})$. We now proceed to the formal proof.

Recall that $H_{d,\bar{d},d}$ is the refinement of $\bar{a}_{d,\bar{d}}$ and $q = \lceil \exp(\delta^{-10}) \rceil$ in dimension $d'$. For every $d, d' \in \{1, 2\}$, $\bar{t} \in T_d$ and $j \in [q]$ it holds that
\begin{align*}
\sum_{i \in H_{d,\bar{d},d'} \cap Q} \bar{a}_{d,\bar{d}} \cdot v_d(i) &\leq \beta \cdot \bar{w} \left(1_{H_{d,\bar{d},d'} \cap \bar{a}_{d,\bar{d}} \cdot \bar{v}_d} + \frac{1}{\varphi^{10}(\delta)} \cdot \text{OPT} \cdot \max \left\{ \sum_{i \in C} 1_{H_{d,\bar{d},d'} \cap \bar{a}_{d,\bar{d}} \cdot \bar{v}_d(i) \mid C \subseteq C \right\} \right) \\
&\leq \beta \cdot \|1_{H_{d,\bar{d},d'} \cap \bar{a}_{d,\bar{d}} \cdot \bar{v}_d} \| + \frac{1}{\varphi^{10}(\delta)} \cdot \text{OPT} \cdot \max \left\{ v_d(H_{d,\bar{d},d'} \cap C) \mid C \subseteq C \right\}.
\end{align*}
The equality follows from the definition of $Q$. The first inequality follows from (29) and $1_{H_{d,\bar{d},d'} \cap \bar{a}_{d,\bar{d}} \cdot \bar{v}_d} \in S_{\text{small}} \subseteq S$. The second inequality holds as $\bar{w}$ is small items integral and $\text{supp}(\bar{a}_{d,\bar{d}}) \subseteq \text{supp}(\bar{w}) \setminus L$.

Thus, by Lemma 4.3 for every $d, d' \in \{1, 2\}$, $\bar{t} \in T_d$ and $j \in [q]$ there is a set $X_{d,\bar{d},d'} \subseteq Q$ such that
\begin{equation}
\left\| 1_{X_{d,\bar{d},d'} \cap \bar{a}_{d,\bar{d}} \cdot (\bar{v}^1 + \bar{v}^2)} \right\| \leq \frac{16}{q} \cdot \text{OPT} + 2q\delta \quad \text{and} \quad \left\| 1_{Q \setminus X_{d,\bar{d},d'} \cap \bar{a}_{d,\bar{d}} \cdot \bar{v}_d} \right\| \leq \beta \cdot \bar{a}_{d,\bar{d}} \cdot \bar{v}_d^d.
\end{equation}
Define $\bar{\mu} \in [0, 1]^{Q \times B}$ by
\begin{equation}
\bar{\mu}_{i,(d,\bar{t},m)} = \begin{cases}
\frac{a_{d,\bar{t}}}{\kappa_d(\bar{t})} & i \in Q \setminus X_{d,\bar{d},1} \setminus X_{d,\bar{d},2} \\
0 & \text{otherwise}
\end{cases}
\end{equation}
for every $i \in Q$, $d \in \{1, 2\}$, $t \in T_d$ and $m \in [\kappa_d(t)]$. Also, for every $i \in Q$ and $m \in [M]$ define

$$\mu_{i,m} = \sum_{d \in \{1, 2\}} \sum_{t \in T_d} \frac{a^{d,\bar{t}}_i \cdot 1_{i \in X^{d,\bar{t},1} \cup X^{d,\bar{t},2}}}{M}.$$ 

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

$$\sum_{b \in B} \mu_{i,b} = \sum_{d \in \{1, 2\}} \sum_{t \in T_d} \sum_{m \in [\kappa_d(t)]} \mu_{i,(d,t,m)} + \sum_{m \in [M]} \bar{\mu}_{i,m} = \sum_{d \in \{1, 2\}} \sum_{t \in T_d} \begin{cases} \kappa_d(t) \cdot a^{d,\bar{t}}_i \cdot \frac{s_d(t)}{M} & i \in Q \setminus X^{d,\bar{t},1} \setminus X^{d,\bar{t},2} \\ M \cdot a^{d,\bar{t}}_i & i \in X^{d,\bar{t},1} \cup X^{d,\bar{t},2} \end{cases} = \sum_{d \in \{1, 2\}} \sum_{t \in T_d} \frac{a^{d,\bar{t}}_i}{M} \cdot \bar{w}_i^1 + \bar{w}_i^2 = 1,$$

where the fourth inequality follows from (27).

For every $d, d' \in \{1, 2\}$, $t \in T_d$ we have

$$\bar{a}^{d,\bar{t}} \cdot \bar{v}_d' = \sum_{i \in I \setminus L} v_d(i) \cdot \bar{a}^{d,\bar{t}}_i \cdot C(i) = \sum_{C \in C^* \text{ s.t. } T^d(C) = \bar{t}} \bar{a}^{d,\bar{t}}_i \cdot \sum_{i \in I \setminus L} v_d(i) \cdot C(i) \leq \sum_{C \in C^* \text{ s.t. } T^d(C) = \bar{t}} \bar{a}^{d,\bar{t}}_i \cdot \left(1 - s_d'(ar{t})\right) \leq \left(1 - s_d'(ar{t})\right) \cdot \eta_d(ar{t}),$$

where the first equality is by (27) and the inequality is by Lemma 4.5. Thus, for every $m \in [\kappa_d(t)]$ we have

$$\sum_{i \in Q \setminus X^{d,\bar{t},1} \setminus X^{d,\bar{t},2}} v_d(i) \cdot \bar{a}^{d,\bar{t}}_i \leq \beta \cdot \frac{\bar{a}^{d,\bar{t}}_i \cdot \bar{v}_d'}{\kappa_d(t)} \leq \frac{\beta \cdot \left(1 - s_d'(ar{t})\right)}{\kappa_d(t)} \leq 1 - s_d'(ar{t}),$$

where the first inequality is by (12) and the last inequality holds since $|T_d| \leq \exp(\delta^{-6})$, $q \geq \exp(\delta^{-10})$ and $M = \exp(-\delta^{-9}) \cdot \text{OPT} + \exp(\delta^{-11})$. Thus, $\bar{\mu} \in P$, i.e., $P \neq \emptyset$.

We now have the tools to prove the following.

**Lemma 4.10.** $\text{OPT}_f(\bar{y}) \leq (1 + 8\delta)|B| + 1$

**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 4.8 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 First-Fit (Lemma 2.7). Thus, $\text{OPT}_f(1_{Q \setminus Q_I}) \leq 8\delta|B| + 1$.

For every $b \in B$ define $C_b = \{i \in Q \mid \bar{\mu}_{i,b} = 1\}$. It follows that $Q_I = \bigcup_{b \in B} C_b$. Recall that $\bar{a}^{d,\bar{t}}$ are the vectors defined in Lemma 4.6. For every $(d, t, m) \in B \setminus [M]$ define a vector $\gamma^{d, t, m} \in [0, 1]^C$ by $\gamma^{d, t, m} = a^{d, \bar{t}} \cdot \bar{C}$ for any $C \in \text{supp}(a^{d, \bar{t}})$, and $\gamma^{d, t, m} = 0$ for any other configuration $C' \in C$. By the definition of $P$, it holds that $v(C_{d, t, m}) \leq 1 - s^{d}(\bar{t})$, and by Lemma 4.6 for every $C \in \text{supp}(a^{d, \bar{t}})$ it holds that $v(C) \leq s^{d}(\bar{t})$; thus, $C \cup C_{d, t, m} \in C$, and $\gamma^{d, t, m}$ is well defined. Also, for any $m \in [M]$ define $\gamma^m \in [0, 1]^C$ by $\gamma^m_{C_m} = 1$ and $\gamma^m_{C_m} = 0$ for $C \in C \setminus \{C_m\}$.
Define $\bar{x} = \sum_{b \in B} \bar{z}^b$. We show that $\bar{x}$ is a solution for \( \text{LP} \left( \sum_{d \in \{1,2\}} \left( (\bar{z}^d - \bar{r}^d) \land I_{L \setminus F_d} \right) + 1_{Q_t} \right) \). For $i \in L$ we have

\[
\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = \sum_{b \in B} \gamma_C^b \cdot C(i) = \sum_{d \in \{1,2\}} \sum_{i \in \mathcal{T}_d} \sum_{m \in [\kappa_d(i)]} \bar{x}_C^{d,I} \cdot C(i) = \sum_{d \in \{1,2\}} \sum_{i \in \mathcal{T}_d} (\bar{z}^d - \bar{r}^d) \land I_{L \setminus F_d}.
\]

The second equality holds by the definition of $\bar{z}^b$ and since the sets $C_b$ do not contain large items. The last equality is by Lemma 4.6. For any $i \in L$, it follows that $\|\bar{x}\| = B$. Thus,

\[
\text{OPT}_f \left( \sum_{d \in \{1,2\}} (\bar{z}^d - \bar{r}^d) \land I_{L \setminus F_d} \right) \leq \|\bar{x}\| = B,
\]

and by the definition of $\bar{y}$ we have

\[
\text{OPT}_f(\bar{y}) = \text{OPT}_f \left( \sum_{d \in \{1,2\}} (\bar{z}^d - \bar{r}^d) \land I_{L \setminus F_d} \right) + \text{OPT}_f(1_{Q_t}) \leq (1 + 8\delta)|B| + 1.
\]

Observe that

\[
|B| = \sum_{d \in \{1,2\}} \sum_{i \in \mathcal{T}_d} \kappa_d(i) + M = \sum_{d \in \{1,2\}} \sum_{i \in \mathcal{T}_d} \left( \lfloor \beta \eta_d(i) \rfloor + 2\delta^{-1} \right) + \exp(-\delta^{-9}) \cdot \text{OPT} + \exp(\delta^{-11})
\]

\[
\leq \beta \|\bar{\lambda}\| + (|T_1| + |T_2|) \cdot (1 + 2\delta^{-1}) + \exp(-\delta^{-9}) \cdot \text{OPT} + \exp(\delta^{-11})
\]

\[
\leq \beta \|\bar{\lambda}\| + \exp(-\delta^{-9}) \cdot \text{OPT} + \exp(\delta^{-12}).
\]

The first inequality holds since $\sum_{i \in \mathcal{T}_d} \eta_d(i) = \|\bar{\lambda}\|$, and the second inequality uses $|T_d| \leq \exp(\delta^{-6})$. By the above, Lemma 4.10 and 4.6, we have

\[
\text{OPT}_f(\bar{z}) \leq (1 + 8\delta)|B| + 1 + 2\beta \delta \|\bar{\lambda}\| + 2\delta^{-5} + \frac{2\delta^{-11}}{\varphi^{10}(\delta)} \text{OPT}
\]

\[
\leq (1 + 8\delta) \left( \beta \|\bar{\lambda}\| + \exp(-\delta^{-9}) \cdot \text{OPT} + \exp(\delta^{-12}) \right) + 1 + 2\beta \delta \|\bar{\lambda}\| + 2\delta^{-5} + \frac{2 \cdot \delta^{-11}}{\varphi^{10}(\delta)} \text{OPT}
\]

\[
\leq \beta(1 + 10\delta) \|\bar{\lambda}\| + \exp(\delta^{-20}) + \delta^{10} \text{OPT},
\]

where the last inequality uses $\varphi = \exp(\delta^{-20})$. Thus, we showed that $\mathcal{S}$ is a linear structure, which completes the proof of Lemma 2.2.

4.3 Refinement for Small Items

Proof of Lemma 4.3: Define $r(i) = \frac{\nu_d(i)}{q_d(i)}$ for any $i \in I$. W.l.o.g assume that $I \setminus L = \{1,2,\ldots,s\} = [s]$ for some $s \in \mathbb{N}$, and $r(1) \leq r(2) \leq \ldots \leq r(s)$. If $\bar{a} \cdot (\bar{v}^1 + \bar{v}^2) \leq \frac{1}{q^2} \text{OPT} + 2q\delta$ define $H_i = I \setminus L$ and $H_j = \emptyset$ for $j \in \{2,\ldots,q\}$. Let $Q \subseteq I \setminus L$ and $\frac{1}{q} \leq \beta \leq 1$ which satisfies (28). We can select $X = I \setminus L$. It follows that $\|1_{Q \setminus X} \cdot \bar{a} \cdot \bar{v}^d\| = 0 \leq \beta \cdot \bar{a} \cdot \bar{v}^d$ and $\|1_X \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\| = \bar{a} \cdot (\bar{v}^1 + \bar{v}^2) \leq \frac{1}{q^2} \text{OPT} + 2q\delta$. This shows the lemma holds in case $\bar{a} \cdot (\bar{v}^1 + \bar{v}^2) \leq \frac{1}{q^2} \text{OPT} + 2q\delta$. We henceforth assume that

\[
\bar{a} \cdot (\bar{v}^1 + \bar{v}^2) > \frac{1}{q^2} \text{OPT} + 2q\delta.
\]
Define $h_0 = 0$ and

$$h_j = \min \left\{ i \in [s] \left| (\bar{a} \wedge 1_{[d]}) \cdot (\bar{v}^1 + \bar{v}^2) \geq \frac{j}{q} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2) \right. \right\} \quad \forall j \in [q] \quad (44)$$

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

Let $Q \subseteq I \setminus L$ and $\frac{1}{q} \leq \beta \leq 1$ which satisfy (28). For every $j \in [q]$ and $C \subseteq C$ it holds that $\nu_d(C \cap H_j) \leq 1$ and

$$\nu_d(C \cap H_j) = \sum_{i \in C \cap H_j} \nu_d(i) = \sum_{i \in C \cap H_j} \nu_d(i) \cdot r(i) \leq r(h_j) \sum_{i \in C \cap H_j} \nu_d(i) \leq r(h_j).$$

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

$$\forall j \in [q]: \max \{ \nu_d(C \cap H_j) \mid C \subseteq C \} \leq \min\{1, r(h_j)\}. \quad (45)$$

We use in our proof the following inequality (that we prove later).

$$\forall j \in [q] \setminus \{1\}: \quad \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| \geq \frac{1}{2} \min\{1, r(h_{j-1})\} \cdot \frac{1}{q^5} \text{OPT}, \quad (46)$$

For every $j \in [q]$ define

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

It follows from (28) and (15) that

$$\beta_j \leq \frac{\text{OPT}}{q^5} \cdot \max \{ \nu_d(C \cap H_j) \mid C \subseteq C \} \leq \min\{r(h_j), 1\} \cdot \frac{\text{OPT}}{q^5}.$$ 

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

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

where the last inequality follows from $\beta \geq \frac{1}{q}$ and (46). Hence, there exists $X_j \neq \emptyset$. As the set is minimal, it follows that there is $x_j \in X_j$ such that $\|1_{X_j \setminus \{x_j\}} \cdot \bar{a} \cdot \bar{v}^d\| \leq \beta_j - 1 \leq \frac{\text{OPT}}{q^5}$. Thus,

$$\|1_{X_j \setminus \{x_j\}} \cdot \bar{a} \cdot \bar{v}^d\| = \sum_{i \in X_j \setminus \{x_j\}} \bar{a}_i \cdot \nu_d(i) = \sum_{i \in X_j \setminus \{x_j\}} \bar{a}_i \cdot \frac{\nu_d(i)}{r(i)} \leq \sum_{i \in X_j \setminus \{x_j\}} \bar{a}_i \cdot \frac{\nu_d(i)}{r(h_{j-1})} \leq \frac{\beta_j - 1}{r(h_{j-1})} \leq \frac{1}{r(h_{j-1})} \min\{\tau_{j-1}, 1\} \cdot \frac{\text{OPT}(I, v)}{q^5} \leq \frac{\text{OPT}(I, v)}{q^5},$$

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

$$\|1_{Q \setminus X} \cdot \bar{a} \cdot \bar{v}^d\| = \sum_{j=1}^{q-1} \|1_{(Q \setminus X) \cap H_j} \cdot \bar{a} \cdot \bar{v}^d\|$$

$$= \|1_{(Q \setminus X) \cap H_1} \cdot \bar{a} \cdot \bar{v}^d\| - \beta_1 + \sum_{j=2}^{q-1} \left( \|1_{(Q \setminus X) \cap H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \beta_j - \beta_j \right) + \beta_{q-1}$$

$$\leq \beta \sum_{j=1}^{q-1} \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \beta_{q-1}$$

$$\leq \beta \sum_{j=1}^{q-1} \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \min\{r(h_{q-1}), 1\} \cdot \frac{\text{OPT}(I, v)}{q^5} \leq \beta \sum_{j=1}^{q} \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| = \beta \cdot \bar{a} \cdot \bar{v}^d.$$
The first equality holds as supp(\(a\)) \(\subseteq \bigcup_{j \in [q]} H_j\). The first inequality follows from the definitions of \(\beta_1\) and \(X_j\) (for \(j \in \{2, \ldots, q-1\}\)). The last inequality follows from \(\beta \geq \frac{1}{q}\) and (46).

Note that \(\|1_{H_q} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\| \leq \frac{\bar{a} \cdot \|v\|}{q} \leq \frac{2 \cdot \text{OPT}}{q}\). Thus,

\[
\|1_X \cdot 1_A \cdot (\bar{v}^1 + \bar{v}^2)\| \leq \|1_{H_q} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\| + \sum_{j=2}^{q} \|1_{X_j} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\|
\leq \frac{2 \cdot \text{OPT}}{q} + q \cdot 2 \cdot \frac{\text{OPT}}{q^2} + 2 \delta q \leq \frac{16}{q} \text{OPT} + 2 \delta q.
\]

It remains to show that (46) holds. For every \(j \in [q]\), we have

\[
\|1_{H_j} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\| = \|1_{h_j} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\| - \|1_{h_{j-1}} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\|
\geq \frac{j}{q} \bar{a} \cdot (\bar{v}^1 + \bar{v}^2) - \frac{j-1}{q} \bar{a} \cdot (\bar{v}^1 + \bar{v}^2) - 2 \delta
= \frac{1}{q} \bar{a} \cdot (\bar{v}^1 + \bar{v}^2) - 2 \delta
\geq \frac{1}{q^2} \text{OPT} + 2 \delta q - 2 \delta
= \frac{1}{q^2} \text{OPT}(I, \nu).
\]

The first inequality follows from (44) and \(v_1(i) + v_2(i) \leq 2 \delta\) for all \(i \in I \setminus L\). The second inequality follows from (43). Additionally, for \(j \in [q] \setminus \{1\}\) we have

\[
\|1_{H_j} \cdot \bar{a} \cdot (\bar{v}^1 + \bar{v}^2)\| = \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\|
= \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \sum_{i \in H_j} \bar{a}_i \cdot v_d(i)
= \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \sum_{i \in H_j} \bar{a}_i \cdot \frac{v_d(i)}{r(i)}
\leq \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| + \sum_{i \in H_j} \bar{a}_i \cdot \frac{v_d(i)}{r(h_{j-1})}
= \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| \cdot \left(1 + \frac{1}{r(h_{j-1})}\right),
\]

where the inequality follows from \(r(1) \leq r(2) \leq \ldots \leq r(p)\). Using (47) and (48), we get

\[
\forall j \in [q] \setminus \{1\}: \quad \|1_{H_j} \cdot \bar{a} \cdot \bar{v}^d\| \geq \left(1 + \frac{1}{r(h_{j-1})}\right)^{-1} \cdot \frac{1}{q^2} \text{OPT} \geq \frac{1}{2} \min\{1, r_{j-1}\} \cdot \frac{1}{q^3} \text{OPT},
\]

where the inequality follows from \((1 + x^{-1})^{-1} \geq \frac{1}{2} \min\{1, x\}\) for every \(x \geq 0\). Inequality (46) follows from the last inequality.

5 Existence of \(\psi\)-Relaxations

In this section we prove Lemmas 2.4, 2.5 and 2.6. That is, we show how to obtain relaxations for various configurations.

Proof of Lemma 2.4: Let \(S \subseteq C \setminus L\) be a minimal set such that either \(v_1(C \setminus S) \leq 1 - \delta\) or \(v_2(C \setminus S) \leq 1 - \delta\). That is, for any \(i \in S\) it holds that \(v(C \setminus (S \setminus \{i\})) > (1 - \delta, 1 - \delta)\). Such a set exists since \(C \in C_0\).

In the following we show that \(v(S) \leq (2 \delta, 2 \delta)\). Assume, by way of contradiction, that \(v_1(S) > 2 \delta\) or \(v_2(S) > 2 \delta\). Then \(S \neq \emptyset\) and there is \(i \in S\). W.l.o.g assume \(v_1(S) > 2 \delta\). Then \(v_1(S \setminus \{i\}) > \delta\) as all the items in \(S\) are small, and \(i \in S\). Therefore,

\[
v_1(C \setminus (S \setminus \{i\})) = v_1(C) - v_1(S \setminus \{i\}) \leq 1 - \delta,
\]
contradicting the definition of $S$. Thus, $v(S) \leq (2\delta, 2\delta)$.

Define $C_1 = C \setminus S$ and $C_2 \in C^*$ 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 $\delta$-slack by the 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 a $\delta$-slack.

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

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

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

For any $i \in S$ it holds that

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

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

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

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,

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

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. \hfill \Box

**Proof of Lemma 2.25.** 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 $1 \leq \ell \leq h - 1$ and $C_h = C \cap L \setminus \{i_h\}$. It can be easily shown that $C_1, \ldots, C_h$ are configurations. Define $\bar{\lambda} \in [0, 1]^{C^*}$ by

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

For any $1 \leq \ell \leq h$ it holds that $i_{\ell}$ 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 the configurations $C_1, \ldots, C_h$ have $\delta$-slack. Thus, for any $C' \in C^*$ with $\bar{\lambda}_{C'} > 0$ it holds that $C'$ has $\delta$-slack. Hence, $\bar{\lambda}$ has a $\delta$-slack.

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

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

For any $i \in C \setminus L$ it holds that $i \in C_{\ell}$ for all $1 \leq \ell \leq h - 1$; thus,

$$\sum_{C' \in C^*} \bar{\lambda}_{C'} \cdot C'(i) = \sum_{j=1}^{h} \frac{1}{h-1} \cdot C_j(i) = \sum_{j=1}^{h-1} \frac{1}{h-1} = 1.$$
For any \( i \in I \setminus C \) we have \( i \notin C_\ell \) for all \( \ell \in [h] \). Therefore,
\[
\sum_{C' \in C^*} \bar{\lambda}_{C'} \cdot C'(i) = \sum_{j=1}^{h} \frac{1}{h-1} \cdot C_j(i) = 0.
\]

Finally,
\[
\| \bar{\lambda} \| = \sum_{C' \in C^*} \bar{\lambda}_{C'} = \sum_{\ell=1}^{h} \bar{\lambda}_{C_\ell} = \frac{h}{h-1}.
\]

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

**Proof of Lemma 2.6**
Define \( C' \in C^* \) by \( C'(i) = \begin{cases} \kappa & i \in C \\ 0 & \text{otherwise} \end{cases} \) where \( \kappa = \lceil \frac{1}{2} \delta^{-1} \rceil \) and \( \bar{\lambda} \in [0,1]^{|C^*|} \)
by \( \bar{\lambda}_{C'} = \frac{1}{\kappa} \) and \( \bar{\lambda}_D = 0 \) for any \( D \in C^* \setminus \{C'\} \). Observe that
\[
v_1(C') = \sum_{i \in I} v_1(i) \cdot C'(i) = \kappa \cdot v_1(C) \leq \left( \frac{1}{2} \delta^{-1} \right) \cdot \delta \leq \left( \frac{1}{2} \cdot \delta^{-1} + 1 \right) \cdot \delta \leq \frac{1}{2} + \delta \leq 0.6 \leq 1 - \delta,
\]
where the last two inequalities follow from \( \delta \in (0,0.1) \). Thus \( C' \) has \( \delta \)-slack and hence \( \bar{\lambda} \) is with \( \delta \)-slack.

For any \( i \in C \) it holds that \( \sum_{D \in C^*} \bar{\lambda}_D \cdot D(i) = \bar{\lambda}_{C'} \cdot C'(i) = \frac{1}{\kappa} \cdot \kappa = 1 \). Also, for any \( i \in I \setminus C \) it holds that \( \sum_{D \in C^*} \lambda_D \cdot D(i) = \lambda_{C'} \cdot C'(i) = 0 \). Finally
\[
\| \bar{\lambda} \| = \frac{1}{\kappa} \leq \frac{1}{\frac{1}{2} \delta} \leq 2 \delta \leq 4 \delta.
\]

Thus \( \bar{\lambda} \) is a \( 4\delta \)-relaxation of \( C \), as required.

### 6 Solving the Matching-LP

In this section we present a PTAS for the MLP problem, thus proving Lemma [13]. 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 \( \delta \)-MLP. Towards this end, we use a result of Grötschel, Lovász, and Schrijver [14] which describes the ellipsoid method via separation oracles. A separation oracle for a polytope \( P \subseteq \mathbb{R}^n \) gets a point \( \bar{x} \in \mathbb{R}^n \) as an input 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, it finds an hyperplane which separates between \( \bar{x} \) and the polytope \( P \). Given a separation oracle, the ellipsoid method either determines that \( P = \emptyset \) or finds \( \bar{x} \in P \) in polynomial time in \( n \). As a consequence, if \( P = \emptyset \) then the execution of the ellipsoid 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 which is commonly used to solve linear programs similar to [11] (see, e.g., [16]). 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 [11] has \(|I|\) variables. This approach cannot be directly implemented for MLP since the number of variables in both primal and dual linear programs is non-polynomial in the \( \delta \)-2VBP instance \((I,v)\), due to the number of linear constraints required to represent the matching polytope. This difficulty is overcome through projections of polytopes in vector space of non-polynomial dimension into polytopes with a polynomial dimension. A similar approach was recently used in [11].

Throughout this section we define multiple mathematical optimization problems. We use \( \text{OPT}(P) \) to denote the value of the optimal solution to the problem \( P \). To avoid notational overhead, we assume the input \( \delta \)-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 \( \delta \)-matching graph of \((I,v)\) as defined in Section [1.2] and \( P_M(G) \) as the matching polytope of \( G \). Recall \( E \) is the projection function defined in Section [1.2].
We first simplify our problem. We can relax the requirement $\sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) = 1$ in (4) and use inequality instead. That is,

$$\text{rMLP : } \min \sum_{C \in \mathcal{C}} \bar{x}_C$$

$$\forall i \in I : \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \geq 1$$

(49)

$$\forall C \in \mathcal{C} : \bar{x}_C \geq 0$$

It can be easily shown that the optimum of (4) and (49) are equivalent, and furthermore, that a solution for (49) can be easily converted to a solution for (4) of the same or lower value.

Our objective is to find a variant of (49) in which the set $\mathcal{C}$ is replaced by a polynomial size set $\mathcal{D} \subseteq \mathcal{C}$, while approximately preserving the optimal value. Towards this end we use the following family of polytopes.

$$\forall \mathcal{D} \subseteq \mathcal{C} : P(\mathcal{D}) = \left\{ \bar{x}, \bar{y} \in \mathbb{R}^n_{\geq 0}, \bar{y} \in P_M(G) \mid \forall i \in I : \sum_{C \in \mathcal{D}} \bar{x}_C \cdot C(i) \geq 1 \right\}$$

(50)

Given $\mathcal{D} \subseteq \mathcal{C}$, with a slight abuse of notation we refer to a vector $\bar{x} \in \mathbb{R}^n_{\geq 0}$ as a vector in $\mathbb{R}^n_{\geq 0}$ where $\bar{x}_C = 0$ for every $C \in \mathcal{C} \setminus \mathcal{D}$. This ensures the term $\mathcal{E}(\bar{x})$ is well defined. Since $P_M(G)$ is downward closed it follows that rMLP is equivalent to the problem of finding $\bar{x}$ such that $||\bar{x}||$ is minimal. For $\mathcal{D} \subseteq \mathcal{C}$ we define the rMLP($\mathcal{D}$) problem as the problem of finding $\bar{x}, \bar{y} \in P(\mathcal{D})$ such that $||\bar{x}||$ is minimal. It follows that $\text{OPT}(\text{rMLP}(\mathcal{D})) \geq \text{OPT}(\text{rMLP})$ for any $\mathcal{D} \subseteq \mathcal{C}$.

We use $P(\mathcal{D})$ to define a family of additional polytopes in $\mathbb{R}^E$.

$$\forall \mathcal{D} \subseteq \mathcal{C}, h \in \mathbb{R}_{\geq 0} : Q(\mathcal{D}, h) = \left\{ \bar{y} \in \mathbb{R}^n \mid \exists \bar{x} \in \mathbb{R}^n_{\geq 0} : (\bar{x}, \bar{y}) \in P(\mathcal{D}) \text{ and } ||\bar{x}|| \leq h \right\}$$

(51)

It thus follows that $Q(\mathcal{D}, h) \neq \emptyset$ if and only if $\text{OPT}(\text{rMLP}(\mathcal{D})) \leq h$. 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 $\bar{y} \in P_M(G)$, and otherwise finds a separating hyperplane using a separation oracle for the matching polytope. If $\bar{y} \in P_M(G)$ we use the following linear program, which depends on $\bar{y} \in P_M(G)$ and $\mathcal{D} \subseteq \mathcal{C}$, to obtain a separating hyperplane.

$$\text{PRIMAL}(\bar{y}, \mathcal{D}) : \min \sum_{C \in \mathcal{D}} \bar{x}_C$$

$$\forall i \in I : \sum_{C \in \mathcal{D}} \bar{x}_C \cdot C(i) \geq 1$$

$$\forall e \in E : \sum_{C \in S(e) \cap \mathcal{D}} \bar{x}_C \leq \bar{y}_e$$

$$\forall C \in \mathcal{C} : \bar{x}_C \geq 0$$

(52)

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 is 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 $\bar{y} \in P_M(G)$ and $\text{OPT}($PRIMAL$(\bar{y}, \mathcal{D})) \leq h$.

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

---

17 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 $\bar{y} \in P$. 

39
Then DUAL($\bar{y}, D$) maximizes \[
\sum_{i \in I} \lambda_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e
\]
subject to:
\[
\forall C \in D \setminus C_2 : \sum_{i \in C} \lambda_i \leq 1
\]
\[
\forall e \in E, \ C \in S(e) \cap D : \sum_{i \in C} \lambda_i \leq 1 + \beta_e
\]
\[
\forall i \in I : \lambda_i \geq 0
\]
\[
\forall e \in E : \bar{\beta}_e \geq 0
\]  

$(53)$

Observe that the feasible region of DUAL($\bar{y}, D$) is independent of $\bar{y}$. That is, for any $D \subseteq C$ we can define

\[
R(D) = \left\{ (\bar{\lambda}, \bar{\beta}) \in \mathbb{R}_0^I \times \mathbb{R}_E^E \left| \begin{array}{l}
\forall C \in D \setminus C_2 : \sum_{i \in C} \bar{\lambda}_i \leq 1 \\
\forall e \in E, \ C \in S(e) \cap D : \sum_{i \in C} \bar{\lambda}_i \leq 1 + \bar{\beta}_e
\end{array} \right. \right\}
\]  

$(54)$

Then DUAL($\bar{y}, D$) is the problem of finding $(\bar{\lambda}, \bar{\beta}) \in R(D)$ for which $\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e$ is maximal.

We use the following relation between $R(D)$ and $Q(D, h)$ to generate separating hyperplanes.

**Lemma 6.1.** For any $h \in \mathbb{R}_{\geq 0}$, $\bar{y} \in Q(C, h)$ and $(\bar{\lambda}, \bar{\beta}) \in R(C)$ 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 $\bar{y} \in Q(C, h)$ it follows that $\text{OPT}(\text{DUAL}(\bar{y}, C)) = \text{OPT}(\text{PRIMAL}(\bar{y}, C)) \leq h$. Thus, as $(\bar{\lambda}, \bar{\beta}) \in R(C)$ we have

\[
\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \leq \text{OPT}(\text{DUAL}(\bar{y}, C)) \leq h.
\]

\[
\square
\]

Let $\mathcal{M}^*$ be a maximum matching in the graph $G$. Since each of the vertices in a matching polytope is a (integral) matching, it holds that

\[
\forall \bar{y} \in P_\mathcal{M}(G) : \sum_{e \in E} \bar{y}_e \leq |\mathcal{M}^*|
\]  

$(55)$

Since for every $e \in \mathcal{M}^*$ it holds that $e \in C_2$ and therefore $v_1(e) > (1 - \delta)$, for every solution $\bar{x}$ of rMLP we have

\[
\sum_{C \in \mathcal{C}} \bar{x}_C \geq \sum_{C \in \mathcal{C}} \bar{x}_C \cdot v_1(C) \geq \sum_{C \in \mathcal{C}} \bar{x}_C \sum_{i \in I} v_1(i) \cdot C(i)
\]

\[
\geq \sum_{i \in I} v_1(i) \sum_{C \in \mathcal{C}} \bar{x}_C \cdot C(i) \geq \sum_{i \in I} v_1(i) \geq \sum_{e \in \mathcal{M}^*} v_1(e) > (1 - \delta)|\mathcal{M}^*|.
\]

Hence

\[
\text{OPT}({\text{rMLP}}) > (1 - \delta)|\mathcal{M}^*|.
\]

We combine Lemma 6.1 with the following lemma, which we prove later in this section.

**Lemma 6.2.** There is a polynomial time algorithm Ellipsoid$_R$ which given $\bar{y} \in P_\mathcal{M}(G)$ and $h > (1 - \delta)|\mathcal{M}^*|$ returns one of the following.

- A subset $D \subseteq C$ such that $|D|$ polynomial in the input size and $\text{OPT}(\text{DUAL}(\bar{y}, D)) \leq (1 + \varepsilon) h$.
- A point $(\bar{\lambda}, \bar{\beta}) \in R(C)$ such that $\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e > h$. 

40
Observe that the execution of the ellipsoid is polynomial. Furthermore, if the algorithm solves \( \text{OPT}(\text{DUAL}(\bar{y})) \) and returns \( \bar{y} \), then by Lemma 6.2 the invocation to Ellipsoid returns a separating hyperplane, thus this is a valid execution of the ellipsoid with a separating hyperplane.

We use the algorithm Ellipsoid from Lemma 6.2 to derive a separation oracle for \( Q(C, h) \). The pseudo code of the oracle is given in Algorithm 3. We note there is a polynomial time separation oracle for the matching polytope (see, e.g, [22]), thus Step 3 can be implemented in polynomial time. While the algorithm does not formally qualify as a separation oracle, it provides the following guarantee.

**Lemma 6.3.** Given \( \bar{y} \in \mathbb{R}^E \) and \( h > (1 - \delta)|M^*| \), Algorithm 3 does one of the following.

- Returns a separating hyperplane between \( Q(C, h) \) and \( \bar{y} \).
- Notifies the ellipsoid to abort and returns \( D \subseteq C \) of polynomial cardinality such that \( \text{OPT}(\text{DUAL}(\bar{y}, D)) \leq (1 + \varepsilon)h \). In this case it must hold that \( \bar{y} \in P_M(G) \).

**Proof.** If \( \bar{y} \notin P_M(G) \) the algorithm finds a separating hyperplane between \( \bar{y} \) and \( P_M(G) \). As \( Q(C, h) \subseteq P_M(G) \) this hyperplane also separates between \( \bar{y} \) and \( Q(C, h) \).

If the invocation to Ellipsoid returns \( (\bar{\lambda}, \bar{\beta}) \in R(C) \) 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 \( Q(C, h) \) by Lemma 6.1. Otherwise, by Lemma 6.2 the invocation to Ellipsoid returns a subset \( D \subseteq C \) of polynomial cardinality such that \( \text{OPT}(\text{DUAL}(\bar{y}, D)) \leq (1 + \varepsilon)h \). It follows that in this case Algorithm 3 notifies the ellipsoid to abort and returns \( D \).

Algorithm 4 utilizes \( \text{Q_separator} \) as a separator oracle. The algorithm may return a vector \( \bar{x} \in \mathbb{R}^D \) for some \( D \subseteq C \). Recall we interpret such a vector as a vector in \( \mathbb{R}^C \) as well.

**Algorithm 4: Ellipsoid**

```
Input: h > (1 - \delta)|M^*|
1 Run the ellipsoid method with \( \text{Q_separator} \) (and \( h \)) as the separation oracle
2 if the ellipsoid returned that the polytope is empty then
3 \quad \text{Return OPT(rMLP)} > h
4 else
5 \quad This case can only happens if the \( \text{Q_separator} \) notified the ellipsoid to abort and returned a set \( D \subseteq C \).
6 \quad Find an optimal solution \( (\bar{x}', \bar{y}') \) for rMLP(D) and return \( \bar{x}' \)
7 end
```

**Lemma 6.4.** Algorithm 4 is a polynomial time algorithm which either determines that \( \text{OPT}(\text{rMLP}) > h \) or finds a solution \( \bar{x}' \) for rMLP with \( ||\bar{x}'|| \leq (1 + \varepsilon)h \).

**Proof.** Observe that the execution of the ellipsoid is polynomial. Furthermore, if the algorithm solves rMLP(D) in Step 5 then by Lemma 6.3 we have that \( |D| \) is polynomial, and hence rMLP(D) can be solved in polynomial time (as there is a separation oracle for \( E(\bar{x}) \in P_M(G) \), and the number of variables and additional constraints is polynomial).

By Lemma 6.3, if the ellipsoid asserts that the polytope is empty it holds that all the invocations to \( \text{Q_separator} \) returned a separating hyperplane, thus this is a valid execution of the ellipsoid with a
separation oracle for $Q(C, h)$. Hence $Q(C, h) = \emptyset$ which implies that $\text{OPT}(rMLP) = \text{OPT}(rMLP(C)) > h$ due to (II).

Otherwise, it must hold that the execution of the ellipsoid method has been aborted by $Q_{\text{separator}}$ at some iteration. Let $\bar{y} \in P_M(G)$ be the value of $\bar{y}$ used in the call to $Q_{\text{separator}}$ in this iteration, let $D \subseteq C$ be the subset of configurations found by $Q_{\text{separator}}$, and let $(\bar{x}', \bar{y}') \in P(D)$ be the solution found in Step 5. It holds that $\|\bar{x}'\| \leq \text{OPT}((\text{PRIMAL}(\bar{y}, D)) = \text{OPT}((\text{DUAL}(\bar{y}, D)) \leq (1 + \varepsilon) h$ where the last inequality is by Lemma 6.3. Since $(\bar{x}', \bar{y}') \in P(D)$ it holds that $E(\bar{x}') \leq \bar{y}' \in P_M(G)$, hence $E(\bar{x}') \in P_M(G)$. From the same reason we also have $\sum_{C \subseteq C} \bar{x}'_C \cdot C(i) \geq 1$ for all $i \in I$. Thus, it holds that $\bar{x}'$ is a solution for rMLP of value at most $(1 + \varepsilon) h$.

\[\square\]

**Algorithm 5: Matching-LP**

1. Run a binary search over the range $(\ell, u) = ((1 - \delta)|M^*|, |I|)$, where in each iteration $\text{Ellipsoid}_{Q}(h)$ is called with $h = \frac{\ell + u}{2}$. If $\text{Ellipsoid}_{Q}$ returned that $\text{OPT}(rMLP) > h$ update $\ell = h$. If $\text{Ellipsoid}_{Q}$ returned a solution $\bar{x}$ update $\bar{x}$ to be the best solution and $u = h$.
2. Repeat the process until $u - \ell < \varepsilon$.
3. If $u \neq |I|$ return the best solution found, otherwise return a vector $\bar{x} \in \{0, 1\}^C$ where $\bar{x}_i = 1$ for every $i \in I$ and $\bar{x}_C = 0$ for any other $C \subseteq C$.

Our algorithm for $\delta$-rMLP, given in Algorithm 6, uses $\text{Ellipsoid}_{Q}$ to guide a binary search.

**Proof of Lemma 6.5.** We show that Algorithm 6 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.

By Lemma 6.4 it holds that $\text{OPT}(rMLP) > \ell$ throughout the binary search, and that 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, it holds that by the end of its execution, the algorithm returns a solution $\bar{x}$ such that

$$\|\bar{x}\| \leq (1 + \varepsilon)|u| < (1 + \varepsilon)(\ell + \varepsilon) \leq \text{OPT}(rMLP) + \varepsilon \leq (1 + 3\varepsilon)\text{OPT}(rMLP),$$

where the last inequality holds since $\text{OPT}(rMLP) \geq 1$ (otherwise $I = \emptyset$ and $\bar{x} = 0$ is an optimal solution).

\[\square\]

We are left to provide the proof for Lemma 6.3. Similarly to $\text{Ellipsoid}_{Q}$, the ellipsoid method is used with an approximate separation oracle. The ellipsoid method is used with polytopes from the following family of polytopes, which we define for every $h \geq 0$, $\bar{y} \in P_M(G)$ and $D \subseteq C$.

$$R(\ell, \bar{y}, D) = \left\{(\bar{\lambda}, \bar{\beta}) \in R(D) \mid \sum_{i \in \ell} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \geq \ell \right\}$$

(56)

To derive a separation oracle for $R(h, \bar{y}, D)$ we use a PTAS for 2-DIMENSIONAL KNAPSACK (2DK) [13]. Using the terminology already defined in this paper, the input for 2DK is a 2VBP instance $(S, v)$, a profits vector $\bar{p} \in \mathbb{R}_{\geq 0}^S$ 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 $p(W) = \sum_{i \in W} \bar{p}_i$ is maximal. We use $(S, v, \bar{p}, \bar{b})$ to denote a 2DK instance. We also allow $\bar{p} \in \mathbb{R}_{\geq 0}^S$ where $S \subseteq T$. The separation oracle is given in Algorithm 5. The pseudo code 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 $\delta$-matching graph $G$.

As in the case of $Q_{\text{separator}}$, it holds that $R_{\text{separator}}$ has properties similar to those of a separation oracle.

**Lemma 6.5.** Algorithm 6 is a polynomial time algorithm which given $(\bar{\lambda}, \bar{\beta}) \in \mathbb{R}^I \times \mathbb{R}^E$, $\bar{y} \in P_M(G)$ and $\ell \geq (1 - \delta)|M^*|$, does one of the following.

...
Returns a separating hyperplane between \( R(\ell, \bar{y}, C) \) and \((\bar{\lambda}, \bar{\beta})\).

Notifies the ellipsoid to abort and returns \((\bar{\lambda}', \bar{\beta}') \in R\left((1 - \frac{\epsilon}{8}) \ell, \bar{y}, C\right)\).

**Proof.** Since there is a PTAS for 2DK \([12]\) it immediately 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 \( R(\ell, \bar{y}, C) \) in \([59]\). Thus, we can assume that \( \sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \geq \ell \) for the remaining part of the proof.

If the algorithm returns a set \( W \) in Step 2 then it holds that \( W \subseteq I \setminus L \) and \( v(W) \leq 1 \) as a solution for 2DK. Thus, \( W \subseteq C \setminus C_2 \) and the inequality \( \sum_{i \in W} \bar{\lambda}_i > 1 \) defines a separating hyperplane due to \([59] \) and \([54] \). Thus, we can assume that the algorithm did not return a set in Step 2 for the remainder of the proof. This implies that the optimal solution for the 2DK instance \((I \setminus L, v, \lambda, 1)\) has value of at most \((1 - \frac{\epsilon}{8})^{-1} \). Since every \( C \subseteq C \) such that \( C \subseteq I \setminus L \) is a solution for \((I \setminus L, v, \lambda, 1)\) it follows that

\[
\forall C \subseteq C, C \subseteq I \setminus L: \sum_{i \in C} \bar{\lambda}_i \leq \left(1 - \frac{\epsilon}{8}\right)^{-1}.
\] (57)

Consider the case in which the algorithm returns the set \( W \cup \{j\} \) in Step 3. It holds that \( v(W \cup \{j\}) \leq v(W) + v(j) \leq 1 - v(\bar{W} \cup \{j\}) = 1 \) as \( W \) is a solution for the 2DK instance \((I \setminus N_G[j], v, \lambda, 1 - v(j))\). Thus \( W \cup \{j\} \subseteq C \). Assume towards a contradiction that \( W \cup \{j\} \subseteq C_2 \). Thus there is \( j' \in W \cap L \) such that \( j', j \) \( \notin \) and we can conclude \( W \cap N[j] \neq \emptyset \), contradiction the definition of \( W \). It therefore holds that \( W \cup \{j\} \subseteq C \setminus C_2 \). Since \( \sum_{i \in W \cup \{j\}} \bar{\lambda}_i > 1 \) the configuration \( W \cup \{j\} \) defines a separating hyperplane due to \([59] \) and \([54] \).

Hence, we can assume that the algorithm did not return a separating hyperplane in Step 3 for the remainder of the proof. Let \( C \subseteq C \setminus C_2 \). If \( C \subseteq I \setminus L \) then it holds that \( \sum_{i \in C} \bar{\lambda}_i \leq (1 - \frac{\epsilon}{8})^{-1} \) by \([57] \). Otherwise there is \( j^* \in C \cap L \). 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 3. It holds that \( C \setminus \{j\} \) is a solution for the 2DK instance \((I \setminus N_G[j], v, \lambda, 1 - v(j))\), thus \( \sum_{i \in C \setminus \{j\}} \bar{\lambda}_i \geq (1 - \frac{\epsilon}{8}) \sum_{i \in C \setminus \{j\}} \bar{\lambda}_i \). Since the algorithm did not return \( W \cup \{j\} \) it also holds that \( \sum_{i \in W \cup \{j\}} \bar{\lambda}_i \leq 1 \). Therefore,

\[
\sum_{i \in C} \bar{\lambda}_i = \bar{\lambda}_j + \sum_{i \in C \setminus \{j\}} \bar{\lambda}_i \leq \bar{\lambda}_j + (1 - \frac{\epsilon}{8})^{-1} \sum_{i \in W} \bar{\lambda}_i \leq \left(1 - \frac{\epsilon}{8}\right)^{-1} \sum_{i \in C} \bar{\lambda}_i \leq \left(1 - \frac{\epsilon}{8}\right)^{-1}.
\]

Thus, we have

\[
\forall C \subseteq C \setminus C_2: \sum_{i \in C} \bar{\lambda}_i \leq \left(1 + \frac{\epsilon}{8}\right)^{-1}.
\] (58)
Next, we consider the case in which the algorithms return a set \(W \cup e\) in Step 7. It holds that 
\[v(W \cup \{e\}) = v(W) + v(e) \leq 1 - v(e) + v(e) = 1\]
since \(W\) is a solution for \((I \setminus L, v, \bar{\lambda}, 1 - v(e))\), hence \(W \cup e \in C\). It follows that \(W \cup \{e\} \in S(e)\). Since \(\sum_{i \in W \cup e} \bar{\lambda}_i > 1 + \beta_e\) it follows that \(W \cup e\) defines a separating hyperplane between \((\bar{\lambda}, \beta)\) and \(R(\ell, \bar{y}, C)\) (due to \(58\) and \(59\)).

We can 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 7 in which \(e = e^*\). It holds that \(C \setminus e \subseteq I \setminus L\) (otherwise it must be that \(v_d(C) > 1\) for some \(d \in \{1, 2\}\)) and \(v(C \setminus e) \leq 1 - v(e)\), thus \(C \setminus e\) is a solution for the 2DK instance \((I \setminus L, v, \bar{\lambda}, 1 - v(e))\). Let \(W\) be the approximate solution found for \((I \setminus L, v, \bar{\lambda}, 1 - v(e))\), it thus holds that \(\sum_{i \in W} \bar{\lambda}_i \geq (1 - \frac{\varepsilon}{8}) \sum_{i \in C \setminus e} \bar{\lambda}_i\). Also, since we assume the algorithm does not return a set in Step 7 it holds that \(\sum_{i \in W \cup e} \bar{\lambda}_i \leq 1 + \beta_e\). Therefore, it holds that

\[
\sum_{i \in C} \bar{\lambda}_i = \sum_{i \in e} \bar{\lambda}_i + \sum_{i \in C \setminus e} \bar{\lambda}_i \leq \sum_{i \in e} \bar{\lambda}_i + \left(1 - \frac{\varepsilon}{8}\right)^{-1} \sum_{i \in W} \bar{\lambda}_i \leq \left(1 - \frac{\varepsilon}{8}\right)^{-1} \left(1 + \beta_e\right).
\]

Let \(e = \{j_1, j_2\}\). Then it holds that \(\{j_1\}, \{j_2\}, C \setminus e \subseteq C \setminus C_2\). Therefore, by \(58\), we have

\[
\sum_{i \in C} \bar{\lambda}_i \leq \bar{\lambda}_{j_1} + \bar{\lambda}_{j_2} + \sum_{i \in C \setminus e} \bar{\lambda}_i \leq 3 \left(1 - \frac{\varepsilon}{8}\right)^{-1}.
\]

The above can be summarized into the following inequality,

\[
\forall e \in E, C \in S(e) : \sum_{i \in C} \bar{\lambda}_i \leq \left(1 - \frac{\varepsilon}{8}\right)^{-1} \left(1 + \min\{\bar{\beta}_e, 2\}\right) = \left(1 - \frac{\varepsilon}{8}\right)^{-1} \left(1 + \bar{\beta}'_e\right).
\]

By \(58\) and \(59\) it holds that \((\bar{\lambda} - \varepsilon/8, \bar{\beta}') \in R(C)\). Furthermore, it holds that

\[
\sum_{i \in I} \left(1 - \frac{\varepsilon}{8}\right) \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e = \left(1 - \frac{\varepsilon}{8}\right) \cdot \left(\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e\right) - \frac{\varepsilon}{4} \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e
\]

\[
\geq \left(1 - \frac{\varepsilon}{8}\right) \cdot \left(\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e\right) - \frac{\varepsilon}{4} \sum_{e \in E} \bar{y}_e
\]

\[
\geq \left(1 - \frac{\varepsilon}{8}\right) \ell - \frac{\varepsilon}{4} \frac{\ell}{1 - \delta}
\]

\[
\geq \left(1 - \varepsilon\right) \ell.
\]

The first inequality holds since \(\bar{\beta}'_e = \min\{\bar{\beta}_e, 2\}\), the second inequality uses \(\sum_{e \in E} \bar{y}_e \leq |M^*| < \frac{\ell}{1 - \delta}\) due to \(58\). Thus, \((\bar{\lambda} - \varepsilon/8, \bar{\beta}') \in R\left((\bar{\lambda} - \varepsilon/8, \ell, \bar{y}, C)\right)\).

Algorithm 7 uses the ellipsoid method with \(R_{\text{separator}}\) as the separation oracle.

**Proof of Lemma 6.2** Note that Ellipsoid\(_R\) runs in polynomial time and furthermore it holds that \(\ell > h > (1 - \delta)|M^*|\), thus \(R_{\text{separator}}\) is used with parameters that match the conditions of Lemma 6.3.

Consider the execution of Algorithm 7. If the ellipsoid returns that the polytope is empty, then all the separation hyperplanes returned by Ellipsoid\(_R\) are also separation hyperplanes with respect to the polytope \(R(\ell, \bar{y}, D)\), thus it must hold that \(R(\ell, \bar{y}, D) = \emptyset\). This implies that \(\text{OPT}(\text{DUAL}(\bar{y}, D)) \leq \ell = \frac{h}{1 - \frac{\varepsilon}{4}} \leq (1 + \varepsilon)h\). Since the execution of the ellipsoid is of polynomial time, it follows that \(|D|\) is also polynomial.

If the ellipsoid was aborted, then by Lemma 6.3 it holds that \((\bar{\lambda}, \bar{\beta}) \in \left((1 - \frac{\varepsilon}{8}, \ell, \bar{y}, C)\right)\). By \(58\) we have that \((\bar{\lambda}, \bar{\beta}) \in R(C)\) and

\[
\sum_{i \in I} \bar{\lambda}_i - \sum_{e \in E} \bar{\beta}_e \cdot \bar{y}_e \geq \left(1 - \frac{\varepsilon}{2}\right) \ell = \left(1 - \frac{\varepsilon}{2}\right) \frac{h}{1 - \frac{\varepsilon}{4}} > h.
\]

\(\square\)
Algorithm 7: Ellipsoid_{R}

Input: $\bar{y} \in P_M(G)$ and $h > (1 - \delta)|M^*|

1. Run the ellipsoid method with $R_{\text{separator}}$ as the separation oracle where $R_{\text{separator}}$ is used with $\bar{y}$ and $\ell = \frac{h}{1 - \frac{\delta}{4}}$.

2. if the ellipsoid returned that the polytope is empty then
   3. Define $D$ as set of configurations which were returned by $R_{\text{separator}}$ as a separating hyperplanes throughout the execution of the ellipsoid. Return $D$.

4. else
   5. This only happens if $R_{\text{separator}}$ aborted the ellipsoid and returned $(\bar{\lambda}, \bar{\beta}) \in ((1 - \frac{\delta}{2}) \ell, \bar{y}, C)$. Return $(\bar{\lambda}, \bar{\beta})$.

6. end

7 References

[1] Nikhil Bansal, Alberto Caprara, and Maxim Sviridenko. A new approximation method for set covering problems, with applications to multidimensional bin packing. *SIAM Journal on Computing*, 39(4):1256–1278, 2010.

[2] Nikhil Bansal, Marek Eliáš, and Arindam Khan. Improved approximation for vector bin packing. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1561–1579. SIAM, 2016.

[3] Chandra Chekuri and Sanjeev Khanna. On multidimensional packing problems. *SIAM journal on computing*, 33(4):837–851, 2004.

[4] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Multi-budgeted matchings and matroid intersection via dependent rounding. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 1080–1097. SIAM, 2011.

[5] Yuan Shih Chow and Henry Teicher. *Probability theory: independence, interchangeability, martingales*. Springer Science & Business Media, 1997.

[6] Henrik I Christensen, Arindam Khan, Sebastian Pokutta, and Prasad Tetali. Approximation and online algorithms for multidimensional bin packing: A survey. *Computer Science Review*, 24:63–79, 2017.

[7] 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.

[8] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. Second edition, 2001.

[9] 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.

[10] Maxence Delorme, Manuel Iori, and Silvano Martello. Bin packing and cutting stock problems: Mathematical models and exact algorithms. *European Journal of Operational Research*, 255(1):1–20, 2016.

[11] Yaron Fairstein, Ariel Kulik, and Hadas Shachnai. Modular and submodular optimization with multiple knapsack constraints via fractional grouping. *CoRR*, abs/2007.10470, 2020.

[12] Yaron Fairstein, Ariel Kulik, and Hadas Shachnai. Modular and submodular optimization with multiple knapsack constraints via fractional grouping. In *29th Annual European Symposium on Algorithms (ESA)*, 2021.
[13] Alan M Frieze and Michael RB Clarke. Approximation algorithms for the m-dimensional 0-1 knapsack problem: worst-case and probabilistic analyses. *European Journal of Operational Research*, 15(1):100–109, 1984.

[14] Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[15] Rebecca Hoberg and Thomas Rothvoss. A logarithmic additive integrality gap for bin packing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2616–2625. SIAM, 2017.

[16] Narendra Karmarkar and Richard M. Karp. An efficient approximation scheme for the one-dimensional bin-packing problem. In *23rd Annual Symposium on Foundations of Computer Science*, pages 312–320. IEEE, 1982.

[17] Colin McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

[18] Rina Panigrahy, Kunal Talwar, Lincoln Uyeda, and Udi Wieder. Heuristics for vector bin packing. *research.microsoft.com*, 2011.

[19] Arka Ray. There is no aptas for 2-dimensional vector bin packing: Revisited. *arXiv preprint arXiv:2104.13362*, 2021.

[20] Sai Sandeep. Almost optimal inapproximability of multidimensional packing problems. *arXiv preprint arXiv:2101.02854*, 2021.

[21] SC Sarin and WE Wilhelm. Prototype models for two-dimensional layout design of robot systems. *IIE transactions*, 16(3):206–215, 1984.

[22] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.

[23] Frits CR Spieksma. A branch-and-bound algorithm for the two-dimensional vector packing problem. *Computers & operations research*, 21(1):19–25, 1994.

[24] Asser N Tantawi and Malgorzata Steinder. Autonomic cloud placement of mixed workload: An adaptive bin packing algorithm. In *2019 IEEE international conference on autonomic computing (ICAC)*, pages 187–193. IEEE, 2019.

[25] Vijay V Vazirani. *Approximation algorithms*. Springer-Verlag, 2001.

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

[27] Yonghong Yu and Yang Gao. Constraint programming-based virtual machines placement algorithm in datacenter. In *International Conference on Intelligent Information Processing*, pages 295–304. Springer, 2012.