RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN
INSTITUT FÜR INFORMATIK DER UNIVERSITÄT BONN

New Results on the Probabilistic Analysis of Online Bin Packing and its Variants

Carsten Oliver Fischer

geboren in Köln

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

Bonn, 2019

1. Gutachter: Prof. Dr. Heiko Röglin

2. Gutachter: Prof. Dr. Thomas Kesselheim

Tag der mündlichen Prüfung: 12. September 2019

Erscheinungsjahr: 2019

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Title page image: Frank Luerweg / Universität Bonn, photography.

Abstract

The *classical bin packing problem* can be stated as follows: We are given a multiset of items $\{a_1, \ldots, a_n\}$ with sizes in $[0, 1]$, and want to pack them into a minimum number of bins, each of which with capacity one.

There are several applications of this problem, for example in the field of logistics: We can interpret the i-th item as time a package deliverer spends for the i-th tour. Package deliverers have a daily restricted working time, and we want to assign the tours such that the number of package deliverers needed is minimized. Another setup is to think of the items as boxes with a standardized basis, but variable height. Then, the goal is to pack these boxes into a container, which is standardized in all three dimensions. Moreover, applications of variants of the classical bin packing problem arise in cloud computing, when we have to store virtual machines on servers.

Besides its practical relevance, the bin packing problem is one of the fundamental problems in theoretical computer science: It was proven many years ago that under standard complexity assumptions it is not possible to compute the value of an optimal packing of the items efficiently – classical bin packing is NP-complete. Computing the value efficiently means that the runtime of the algorithm is bounded polynomially in the number of items we have to pack.

Besides the offline version, where we know all items at the beginning, also the *online version* is of interest: Here, the items are revealed one-by-one and have to be packed into a bin immediately and irrevocably without knowing which and how many items will still arrive in the future. Also this version is of practical relevance. In many situations we do not know the whole input at the beginning: For example we are unaware of the requirements of future virtual machines, which have to be stored, or suddenly some more packages have to be delivered, and some deliverers already started their tour.

We can think of the classical theoretical analysis of an online algorithm A as follows: An adversary studies the behavior of the algorithm and afterwards constructs a sequence of items I. Then, the performance is measured by the number of used bins by A performing on I, divided by the value of an optimal packing of the items in I. The adversary tries to choose a worst-case sequence so this way to measure the performance is very pessimistic. Moreover, the chosen sequences I often turn out to be artificial: For example, in many cases the sizes of the items increase monotonically over time.

Instances in practice are often subject to random influence and therefore it is likely that they are good-natured. In this thesis we analyze the performance of online algorithms with respect to two stochastic models.

1. The first model is the following: The adversary chooses a set of items $\mathcal{I}$ and a distribution F on $\mathcal{I}$. Then, the items are drawn independently and identically distributed according to F.
2. In the second model the adversary chooses a finite set of items $\mathcal{I}$ and then these items arrive in random order, that is random with respect to the uniform distribution on the set of all possible permutations of the items.

It is possible to show that the adversary in the second stochastic model is at least as powerful as in the first one.

We can classify the results in this thesis in three parts:

- In the first part we consider the complexity of classical bin packing and its variants cardinality-constrained and class-constrained bin packing in both stochastic models. That is, we determine if it is possible to construct algorithms that are in expectation nearly optimal for large instances that are constructed according to the stochastic models or if there exist non-trivial lower bounds. Among other things we show that the complexity of class-constrained bin packing differs in the two models under consideration.

- In the second part we deal with *bounded-space* bin packing and the dual maximization variant *bin covering*. We show that it is possible to overcome classical worst-case bounds in both models. In other words, we see that bounded-space algorithms benefit from randomized instances compared to the worst case.
- Finally, we consider selected heuristics for class-constrained bin packing and the corresponding maximization variant class-constrained bin covering. Here, we note that the different complexity of class-constrained bin packing with respect to the studied stochastic models observed in the first part is not only a theoretical phenomenon, but also takes place for many common algorithmic approaches. Interestingly, when we apply the same algorithmic ideas to class-constrained bin covering, we benefit from both types of randomization similarly.

This research was supported by ERC Starting Grant 306465 (BeyondWorstCase).

Zusammenfassung

Wir können das *klassische Bin-Packing Problem* wie folgt formulieren: Als Eingabe ist eine Multimenge von Items $\{a_1, \ldots, a_n\}$ mit Größen in $[0, 1]$ gegeben. Diese Items wollen wir in Behälter packen, die jeweils Kapazität 1 besitzen. Hierbei soll die Zahl der verwendeten Behälter minimiert werden.

Es lassen sich verschiedene Anwendungen dieses Problems zum Beispiel im Bereich der Logistik finden. So können wir die Items als Touren für LKW-Fahrer interpretieren, die jeweils Zeit a_i benötigen. Die Behälter stellen in diesem Fall die Fahrer dar, deren tägliche Fahrzeit eine gewisse Zahl an Stunden nicht überschreiten darf. Ziel ist es dann möglichst wenig Fahrer einsetzen zu müssen. Eine andere Möglichkeit ist es die Items als Boxen mit einer genormten Grundfläche, aber unterschiedlichen Höhen anzusehen, die wir in einen Container packen wollen, dessen Maße in allen drei Dimensionen festgelegt sind. Weiterhin spielen Varianten des Problems im Bereich des Cloud-Computings eine Rolle, beispielsweise um virtuelle Maschinen auf Servern zu verwalten.

Neben der praktischen Relevanz handelt es sich beim klassischen Bin-Packing Problem um eines der fundamentalen Probleme der theoretischen Informatik: Schon früh stellte sich heraus, dass sich eine optimale Packung der Items in Behälter unter Standardannahmen der Komplexitätstheorie nicht effizient berechnen lässt – das Problem ist NP-vollständig. Effizient bedeutet hier, dass die Laufzeit eines entsprechenden Algorithmus polynomiell in der Zahl der zu packenden Items beschränkt ist.

Neben der Offline-Version bei der alle zu packenden Items zu Beginn bekannt sind, ist auch die Online-Version von großem Interesse: Hierbei werden die Items nach und nach aufgedeckt und müssen dann sofort in einen der Behälter gepackt werden – ohne das Wissen welche und wie viele Items überhaupt noch in Zukunft gepackt werden müssen. Ferner dürfen einmal gepackte Items nicht später wieder umgepackt werden. Auch diese Problemstellung ist von praktischer Relevanz. Oft kennen wir nicht die gesamte Eingabe zu Beginn: So wissen wir nicht welche Anforderungen zukünftige virtuelle Maschinen haben, die wir auf unseren Servern verwalten werden, oder es muss überraschenderweise eine weitere Tour von einem Fahrer übernommen werden, die aber teilweise bereits losgefahren sind.

Wir können uns die klassische theoretische Analyse eines Online-Algorithmus A wie folgt vorstellen: Ein Gegner untersucht den Algorithmus und konstruiert dann eine Abfolge von Items I. Die *Performance* von A wird dann ermittelt indem die Zahl der geöffneten Bins bei der Bearbeitung von I durch A durch die Zahl der geöffneten Bins in einer optimalen Packung der Items in I geteilt wird.

Da der Gegner bestrebt ist eine möglichst schlechte Sequenz zu erzeugen, ist diese Art der Analyse eines Algorithmus sehr pessimistisch und die konstruierten Folgen von Items I wirken oft künstlich: Beispielsweise steigt die Größe der Items häufig monoton an. Die in der Praxis auftretenden Instanzen unterliegen oftmals stochastischen Einflüssen, und daher hegen wir die Hoffnung, dass diese gutartiger sind.

In dieser Arbeit wollen wir die Performance von Online-Algorithmen für verschiedene Varianten des Bin-Packing Problems bezüglich zweier stochastischer Modelle untersuchen:

1. Im ersten Modell darf der Gegner eine Menge von Items $\mathcal{I}$ mit einer Verteilung F auf diesen vorgeben. Die Items werden dann unabhängig und identisch verteilt bezüglich F generiert.
2. Im zweiten Modell darf der Gegner eine endliche Menge von Items $\mathcal{I}$ vorgeben, die dann in einer zufälligen Reihenfolge aufgedeckt werden. Zufällige Reihenfolge bedeutet hier, dass wir eine Permutation der Items bezüglich der Gleichverteilung auf allen Permutationen ziehen.

Man kann zeigen, dass das zweite Modell mindestens so mächtig ist wie das erste. Das heißt, dass es dem Gegner eher möglich ist Instanzen zu erzeugen auf denen die Performance der Algorithmen schlecht ist.

Die Ergebnisse der Dissertation können in drei Teile gegliedert werden:

- Im ersten Teil beschäftigen wir uns mit der grundsätzlichen Komplexität des klassischen Bin-Packing Problems und der Varianten cardinality-constrained und class-constrained Bin-Packing in den beiden stochastischen Modellen. Konkret setzen wir uns mit der Frage auseinander ob es möglich ist Algorithmen zu konstruieren, die im Erwartungswert auf sehr großen Instanzen, die gemäß der beiden Modelle erzeugt werden, nahezu optimal sind, oder ob nicht-triviale untere Schranken existieren. Unter anderem zeigen wir, dass die Variante *class-constrained* Bin-Packing in den beiden stochastischen Modellen eine unterschiedliche Komplexität aufweist.
- Im zweiten Teil beschäftigen wir uns mit der Variante des Problems, in dem die Algorithmen nur auf eine beschränkte Anzahl an Behältern gleichzeitig zugreifen können, sogenanntes *bounded-space* Bin-Packing. Hier zeigen wir, dass es möglich ist bekannte Schranken, die im Worst-Case gelten, in beiden Modellen zu überwinden. Dies bedeutet, dass solche Algorithmen gegenüber dem Worst-Case von zufällig erzeugten Instanzen profitieren.
- Schließlich untersuchen wir ausgewählte Heuristiken für die Variante des class-constrained Bin-Packings und für die duale Version des Problems, nämlich class-constrained Bin-Covering. Hierbei stellen wir fest, dass das im ersten Teil beobachtete Phänomen, dass class-constrained Bin-Packing in den beiden betrachteten stochastischen Modellen unterschiedlich schwierig ist, nicht nur theoretischer Natur ist: Für zahlreiche algorithmische Ansätze stellt sich heraus, dass sich diese im zweiten Modell schlechter verhalten als im ersten. Interessanterweise stellen wir diesen Effekt nicht fest, wenn wir das duale Problem class-constrained Bin-Covering betrachten: Dort profitieren diese Ansätze gleichermaßen vom randomisierten Input.

Danksagung

Die letzten fünf Jahre, die ich mit den Mitgliedern der Abteilungen V und I verbracht habe, stellten für mich eine wunderbare Zeit dar. Ich habe mich jeden Tag gefreut ins Büro gehen zu können, habe dies nie als Arbeit empfunden, und so lassen sich sicherlich auch einige verfallene Urlaubstage erklären.

Zuallererst möchte ich mich bei Heiko Röglin dafür bedanken, dass er es mir ermöglichte diese fünf Jahre als Doktorand in der theoretischen Informatik an der Universität Bonn zu verbringen, mich dabei beständig unterstützte und jederzeit ein offenes Ohr hatte.

Während meiner Promotion habe ich das Büro mit vielen verschiedenen Mitgliedern dieser Abteilungen geteilt: Angefangen mit Simone und Shoshanna, später dann für lange Zeit mit Barbara und David und schließlich mit Xianghui. Ihnen gilt mein Dank für das immer harmonische Miteinander.

Abgesehen von meinen Bürokollegen, möchte ich mich bei Alex, Andreas, Clemens, Melanie und Michael für sehr viele Diskussionen und Gespräche bedanken. Diese Unterhaltungen waren fachlich sehr hilfreich und ich habe viel durch diese gelernt. Aber auch abseits des fachlichen haben sie mir viele neue Horizonte eröffnet und ich werde diese in Zukunft sicherlich vermissen.

Da sich nun – nach vielen Jahren – auch meine Zeit an der Universität Bonn dem Ende zuneigt, möchte ich mich an dieser Stelle auch bei den Freunden bedanken, die ich im Laufe des Studiums gewonnen habe. Hierbei sind sicherlich insbesondere Lobby, Martin, die beiden Raphaels und Thomas zu nennen. Ohne eure Unterstützung und Hilfsbereitschaft während des Studiums wäre es sicherlich nicht zu dieser Arbeit gekommen.

Schließlich gilt mein Dank meinen Eltern sowie meinem Bruder Christoph.

Contents

1 Introduction

Today's economy would not be possible without modern logistics. This becomes clear looking at e-commerce: Many products we buy now via the Internet, were sold through department stores only a few years ago. This has drastic consequences for the transport of goods: Nowadays it is necessary to deliver many different products to customers, who are distributed all over the city. In general, globalization has made production chains more complex and spread over different continents. This is only possible if transport and warehousing are perfectly organized.

Apart from the mentioned changes, the area of logistics will be subject to drastic transformations in the future: More and more aspects like sustainability and reduction of greenhouse gases will play a role. Furthermore, new technologies like Big Data, Machine Learning and Artificial Intelligence as well as drones and autonomous driving will become important.

From the view point of a theoretical computer scientist the field of logistics is interesting as many algorithmic problems like Shortest-Path-, the Travelling-Salesman- or the Facility-Location-Problem naturally arise. Moreover, many new algorithmic problems are inspired by real-world problems. One of the fundamental problems in theoretical computer science is the *classical bin packing problem.* This problem is as follows:

Classical Bin Packing
Input: A multiset of items $\mathcal{I} = \{a_1, \ldots, a_n\}$ with $a_i \in [0, 1]$.
Task: Find a minimum number ℓ of disjoint sets $B_1, \ldots, B_\ell \subseteq \mathcal{I}$ with $\bigcup_{i=1}^{\ell} B_i = \mathcal{I}$ such that $\sum_{a \in B_j} a \leq 1$ for all $j \in \{1, \ldots, \ell\}$.

The disjoint sets can be interpreted as bins that have a capacity of one.

Also, for classical bin packing we can think of different applications in logistics: The item sizes can represent the length of tours package deliverers have to perform, where the daily working hours of the drivers are limited. Another possibility is to interpret the items as boxes with a standardized basis, but arbitrary height, which we want to pack into a container which is standardized in all three dimensions. Moreover, similar problems arise in cloud computing where virtual machines have to be stored on servers.

When we want to solve a problem using an algorithm we are usually interested in two questions:

- What is the quality of the solution the algorithm computes?
- What is the runtime of the algorithm?

At first, we want to consider the question of how to measure the performance of an algorithm A for classical bin packing: It was already mentioned that the disjoint sets can be interpreted as bins. We think of the algorithm opening these bins and packing items into them. Measuring the performance of A can be seen as a game between the designer of the algorithm and an adversary. The adversary studies the behavior of the algorithm, and then chooses a set of items $\mathcal{I}$ for which A has to compute a packing. Then, the number of bins used by A is divided by the number of opened bins in an optimal packing.

Of course, we want to find algorithms that compute an optimal packing of $\mathcal{I}$ efficiently. Here, efficiently means that the runtime of the algorithm is polynomially bounded in the number of items that have to be packed. Unfortunately, for a large class of problems it remains unclear if it is possible to compute an optimal solution efficiently. The predominant opinion is that it is impossible. This is broadly speaking the $\mathsf{P} \neq \mathsf{NP}$-conjecture – one of the seven Millennium Prize Problems. Besides many other problems that are important in practice, also for classical bin packing it is unclear if we can solve it efficiently.

If it is unclear how to compute an optimal solution efficiently, an usual approach is to find heuristics with better runtime. Such heuristics do not compute necessarily an optimal solution, but a solution which is close.

Let I denote an instance for classical bin packing, and let $\mathrm{OPT}(I)$ denote the number of opened bins in an optimal packing of these items. For classical bin packing it is possible to design for every ϵ greater than zero an algorithm A_ϵ that uses at most $\epsilon\,\mathrm{OPT}(I)$ bins more than an optimal solution for sufficiently large instances. And these algorithms compute the solution efficiently – at least from a theoretical point of view.

Presented in this way classical bin packing is an offline-problem, that is the algorithm knows all items that must be packed before beginning with packing them into bins. Besides offline-problems, also *online-problems* are of great interest: In online-problems the algorithm is *not* aware of all the input at the beginning. The online version of classical bin packing is the following: The items are revealed one by one and every time a new item is revealed the algorithm has to pack it into a bin neither knowing the number of items that will arrive in the future nor any other information about them. Moreover, the algorithm is not allowed to repack any packed items.

The performance of an online algorithm A can be measured similarly to the offline case: An adversary chooses an instance I (this time including an ordering of the items) and then the number of opened bins by A is compared to the number of opened bins according to an optimal offline solution $\mathrm{OPT}(I)$. So, also for an online algorithm we compare the computed solution with the corresponding offline solution.

Of course, online-problems are usually harder than offline-problems. So, another question naturally arises: *How well can an online algorithm perform without knowing the future?* When we consider this question we often drop runtime requirements.

Sometimes it is impossible to obtain good results without knowledge of the future. The *online selection problem* is the following: An adversary presents us one by one n real numbers $v_1, \ldots, v_n$: Each time a new value is revealed we have to decide if we accept it and reject all values which will be revealed in the future, or if we reject this value and wait for the next one. Our goal is to choose the greatest v_i. It is clear that the quality of the solution of any deterministic or randomized online algorithm for this problem can be arbitrarily bad for large instances. For online classical bin packing things are better: There exists an algorithm that opens at most $1.5783\,\mathrm{OPT}(I)$ many bins [8]. On the other hand, for every algorithm A, no matter how sophisticated, there exists an instance I such that A must open at least $1.5427\,\mathrm{OPT}(I)$ many bins [9]. So, online classical bin packing is in fact more complicated than the corresponding offline variant.

The presented way of measuring the quality of an algorithm is most pessimistic and therefore called worst-case analysis. It often turns out that the instances designed by the adversary are in some sense artificial. For example in online bin packing the item sizes are usually presented in increasing order. The same holds true for the online selection problem.

In practice, it is often unlikely that such artificial instances occur, for example because of stochastic influences. So is it possible to obtain better results if the adversary is weakened? An important variant of the online selection problem allows the adversary to choose the values $v_1, \ldots, v_n$, but then presents them to the algorithm in random order. This is the so-called *secretary problem*. In this setting it is possible to select the best candidate with probability $1/e \approx 0.368$ – a huge improvement in comparison to the worst case [34, 62]. In this thesis, we will investigate how online algorithms for variants of bin packing benefit from a randomized input compared to the worst case.

For offline-problems there is a similar phenomenon: There are algorithms that compute an optimal solution, exhibit an exponential worst-case runtime, but perform well in practice. Here, worst-case instances (this time with respect to the runtime) are sometimes artificial too. To bridge this gap between theory and practice in the offline case models like the smoothed-analysis-model have been introduced [80]: In this model an adversary chooses an instance, but afterwards this instance is slightly perturbed in a random way. Among other things, this can be justified by measurement errors occurring in practice. Then, for some algorithms it was shown that the expected runtime is only polynomial – in contrast to their exponential worst-case runtime [2, 35, 80].

We will study variants of online bin packing, where the input is usually generated according to one of the following two stochastic models:

- In the first model the adversary chooses a set of items $\mathcal{I}$ and a distribution F on this set. Then, the instances are generated by drawing items that are independent and distributed according to F.
- In the second model the adversary chooses a finite set of items $\mathcal{I}$, and afterwards a random permutation of these items is generated and revealed to the algorithm. This model is also called *random-order model.*

In both cases the adversary is weakened in comparison to the worst case. Moreover, it is possible to show that the adversary in the random-order model is at least as powerful as in the first model.

The focus of this thesis will be on the following questions:

- What is the best possible performance an algorithm can obtain for the studied problem variants with respect to the stochastic models; and
- How well do selected heuristics behave in these settings?

The remaining part of this introduction is structured as follows: In Subsection 1.1 we will give a formal definition of all variants of the classical bin packing problem that we will consider and the performance measures that are relevant for this thesis. Then, in Subsection 1.2 we will present common algorithmic approaches for solving these problems. Subsection 1.3 will contain an overview of results that are important in the area of bin packing and finally, in Subsection 1.4 we will give an overview of the results presented in this thesis and will relate them with corresponding results in the literature.

1.1 Bin Packing Variants and Performance Measures

As classical bin packing is a fundamental problem in computer science lots of different variants of this problem are studied. At this point we introduce the variants that are relevant for this thesis.

The following notation will be useful in the future: For an $\ell \in \mathbb{N}$ we set $[\ell] := \{1, \dots, \ell\}$.

Cardinality-constrained Bin Packing
Input: A multiset of items $\mathcal{I} = \{a_1, \dots, a_n\}$ with $a_i \in [0, 1]$ and a parameter $k \in \mathbb{N}$.
Task: Find a minimum number ℓ of disjoint sets $B_1, \dots, B_\ell \subseteq \mathcal{I}$ with $\bigcup_{i=1}^{\ell} B_i = \mathcal{I}$ such that $\sum_{a \in B_j} a \leq 1$ and $|B_j| \leq k$ for all $j \in [\ell]$.

> Class-constrained Bin Packing
> *Input:* A multiset of items $\mathcal{I} = \{(s_1, c_1), \ldots, (s_n, c_n)\}$ with $s_i \in [0,1]$ and $c_i \in \mathbb{N}$, and a parameter $k \in \mathbb{N}$.
> *Task:* Find a minimum number ℓ of disjoint sets $B_1, \ldots, B_\ell \subseteq \mathcal{I}$ with $\bigcup_{i=1}^{\ell} B_i = \mathcal{I}$ such that $\sum_{(s,c)\in B_j} s \le 1$ and $|\{c \in \mathbb{N} : (s,c) \in B_j\}| \le k$ for all $j \in [\ell]$.

Classical, cardinality-constrained and class-constrained bin packing are minimization problems, and sometimes we will subsume them as *packing problems*. It is useful to think of these problems as follows: We are given *bins* with a capacity and want to pack the items into the bins such that the capacity-condition (and the side-condition) is satisfied, and we want to minimize the number of non-empty bins.

Moreover, we will also deal with two corresponding maximization versions:

> Classical Bin Covering
> *Input:* A multiset of items $\mathcal{I} = \{a_1, \ldots, a_n\}$ with $a_i \in [0,1]$.
> *Task:* Find a maximum number ℓ of disjoint sets $B_1, \ldots, B_\ell \subseteq \mathcal{I}$ such that $\sum_{a\in B_j} a \ge 1$ for all $j \in [\ell]$.

> Class-constrained Bin Covering
> *Input:* A multiset of items $\mathcal{I} = \{(s_1, c_1), \ldots, (s_n, c_n)\}$ with $s_i \in [0,1]$ and $c_i \in \mathbb{N}$, and a parameter $k \in \mathbb{N}$.
> *Task:* Find a maximum number ℓ of disjoint sets $B_1, \ldots, B_\ell \subseteq \mathcal{I}$ such that $\sum_{(s,c)\in B_j} s \ge 1$ and $|\{c \in \mathbb{N} : (s,c) \in B_j\}| \ge k$ for all $j \in [\ell]$.

Here, it is useful to interpret the capacity of a bin as demand. We will refer to both problems as *covering problems*.

The second attribute of an item in class-constrained bin packing and bin covering is understood as *color* of the item. Moreover, in the class-constrained versions we will also study the variant of unit sized items: In this case every item has size one and the capacity/demand of a bin is given by a parameter $B \in \mathbb{N}$.

It is clear how the online versions of the described problems look like: An unknown number of items will be revealed one by one and we have to pack them irrevocably into bins without knowledge about possible further arriving items.

We notice that the class-constrained versions generalize both the classical and the cardinality-constrained setups: To obtain classical bin packing or bin covering from the class-constrained version we set k equal to one and equip every item with the same color. If we want to obtain cardinality-constrained bin packing from class-constrained bin packing we equip each item with a different color and choose the same parameter k.

Usually, we will denote an instance for an online problem by I. We can think of I as a vector of items, which encodes the order of the arrival of the items, and the entries are revealed to the algorithm one by one. An offline algorithm is allowed to look at all the items before packing them, so the order of the arrivals is irrelevant. In this case an instance is fully described by the multiset of items $\mathcal{I}$ that have to be packed. But we can also use online algorithms for the offline version by assuming a fixed order of arrival.

Let A be an (online) algorithm for one of the presented packing or covering problems. Then, $A(I)$ denotes the number of opened/covered bins of the solution calculated by A on I. Moreover, $\mathrm{OPT}(I)$ (or $\mathrm{OPT}(\mathcal{I})$) denotes the value of an optimal offline packing of the items in I (or $\mathcal{I}$); that is, the minimum number of bins we have to open to pack all items or the maximum number of bins that can be covered.

Now, we describe how to measure the performance of an algorithm A: The classical way to measure the performance of an offline algorithm in the worst case is the *approximation ratio*:

▶ **Definition 1.** Let P be a packing problem, and let $\mathcal{S}$ denote the set of feasible instances for P. Then, the approximation ratio of a deterministic algorithm A for P is defined as

$$\mathrm{AR}(A) = \limsup_{m\to\infty} \sup_{I\in\mathcal{S}\,:\,\mathrm{OPT}(I)\geq m} \frac{A(I)}{\mathrm{OPT}(I)}.$$

We notice that here the performance is measured in an asymptotic sense. Artificial instances often turn out to be small, so they are ruled out by considering only large instances. But it is also possible to require more: The *absolute* approximation ratio is defined as $\mathrm{AR}_a(A) = \sup_{I\in\mathcal{S}} \frac{A(I)}{\mathrm{OPT}(I)}$.

We point out that in the literature the approximation ratio is also called *asymptotic approximation ratio* and the absolute approximation ratio is only called *approximation ratio*. The reason for our convention is that most of our results concern the asymptotic case.

For online algorithms we measure the performance in the worst case by the *competitive ratio*:

▶ **Definition 2.** Let P be a packing problem, and let $\mathcal{S}$ denote the set of feasible instances for P. Then, the competitive ratio of a deterministic algorithm A for P is defined as

$$\mathrm{CR}(A) = \limsup_{m\to\infty} \sup_{I\in\mathcal{S}\,:\,\mathrm{OPT}(I)\geq m} \frac{A(I)}{\mathrm{OPT}(I)}.$$

We notice that this is exactly the same definition as for the approximation ratio. The difference is that the performance of an *online* algorithm is still compared to an optimal *offline* solution. Again, there is also a stronger measure, the *absolute* competitive ratio: $\mathrm{CR}_a(A) = \sup_{I\in\mathcal{S}} \frac{A(I)}{\mathrm{OPT}(I)}$.

The definitions here are given for a packing problem. In this case we always have $\mathrm{CR}_a(A) \geq \mathrm{CR}(A) \geq 1$. If we deal with covering problems we have to replace $\limsup$ by $\liminf$ and $\sup$ by $\inf$ in the previous definitions. Then, it follows that $0 \leq \mathrm{CR}_a(A) \leq \mathrm{CR}(A) \leq 1$.

In Section 3-5 we investigate online algorithms for packing and covering problems with respect to probabilistic performance measures. A distribution F is a pair of a (possibly uncountable) set of items $\mathcal{I}$ and a probability mass function/probability density function p. For a distribution F and $n \in \mathbb{N}$ we set $I_n^F = (A_1, \ldots, A_n)$, where the random variables A_i are independent and distributed according to F. Now we are able to define the *average performance ratio*:

▶ **Definition 3.** Let P be a packing problem and A a deterministic online algorithm for this problem. For a distribution F we set

$$\mathrm{APR}(A, F) = \limsup_{n\to\infty} \mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right],$$

and for a set of distributions $\mathcal{D}$ we set

$$\mathrm{APR}(A, \mathcal{D}) = \sup_{F\in\mathcal{D}} \mathrm{APR}(A, F).$$

In case of a covering problem we have to replace again $\limsup$ by $\liminf$ and $\sup$ by $\inf$. We notice that in the previous definition the number of items tends to infinity and not the value of an optimal solution. But if the items are sampled independently and identically distributed, then also the expected value of an optimal solution grows linearly.

The second performance measure we consider is the *random-order ratio*. Let $\mathcal{I} = \{a_1, \dots, a_n\}$ be a finite multiset of n items. For a random permutation σ of the elements in $[n]$ we set $A_i = a_{\sigma(i)}$ for $i \in [n]$ and $I^\sigma = (A_1, \dots, A_n)$.

▶ **Definition 4.** Let P be a packing problem and A a deterministic online algorithm for this problem. Then, the random-order ratio is defined as

$$\mathrm{RR}(A) = \limsup_{m\to\infty} \sup_{\mathcal{I}\,:\,\mathrm{OPT}(\mathcal{I})\geq m} \frac{\mathbb{E}\left[A(I^\sigma)\right]}{\mathrm{OPT}(\mathcal{I})},$$

where I^σ is a random permutation of the items in $\mathcal{I}$ (that is with respect to the uniform distribution on the set of all possible permutations).

Again, if we deal with a covering problem instead we have to adjust the definition.

The performance measures were only stated for deterministic algorithms. If we consider randomized algorithms, we look at the *expected* number of opened/covered bins by the algorithm on the (possibly random) instances instead.

For a minimization problem it is clear that we have $\mathrm{CR}(A) \geq \mathrm{RR}(A)$. Moreover, for an arbitrary set $\mathcal{D}$ of distributions we have $\mathrm{RR}(A) \geq \mathrm{APR}(A, \mathcal{D})$. So, the random-order ratio is something in between the worst case and a setting where items are drawn independently and identically distributed. Furthermore, this connection allows us to obtain lower bounds for the random-order ratio by considering the case of independent, identically distributed items. We will prove this relationship in Section 2.4.

1.2 Overview of Algorithms

We start with introducing some fundamental algorithms. Most of these heuristics will be investigated in this thesis with respect to the mentioned probabilistic performance measures. Moreover, some algorithms, which we will design in this thesis, are based on such algorithmic approaches.

We start with the classical bin packing problem:

- Next-Fit (NF): At each point of time one bin is open. NF assigns each arriving item to the currently open bin if it can accommodate the item. Otherwise it closes the currently open bin and opens a new bin to which the item is added. Here closing a bin means that no item will be assigned to this bin in the future anymore.
- First-Fit (FF): FF never closes a bin, i.e., it keeps all bins open and assigns each arriving item to the first bin that can accommodate it if such a bin exists. Otherwise it opens a new bin and adds the item to it.
- Best-Fit (BF): BF never closes a bin, i.e., it keeps all bins open and assigns each arriving item to the fullest bin that can accommodate it if such a bin exists. Otherwise it opens a new bin and adds the item to it.
- Worst-Fit (WF): WF never closes a bin, i.e., it keeps all bins open and assigns each arriving item to the bin with the most space remaining that can accommodate it. If such a bin does not exists it opens a new bin and adds the item to it.
- Smart-Next-Fit (SNF): SNF works similarly to NF. It assigns each arriving item to the currently open bin Z if this bin can accommodate the item. Otherwise it opens a new bin Z' and adds the item to it. It retains as new current bin whichever of Z and Z' has the most space remaining.

- HARMONIC_M: This algorithm partitions the interval $[0, 1]$ into the subintervals $[0, 1/M]$, $(1/M, 1/(M-1)]$, $\dots$, $(1/3, 1/2]$, $(1/2, 1]$. These subintervals are used to classify arriving items according to their size. The algorithm packs items from each subinterval independently, using Next-Fit.

While in NF there is only one and in HARMONIC_M only M open bins at each point of time, in FF, BF, and WF all bins are kept open during the whole input sequence. We say that an algorithm is an *ℓ-bounded space algorithm* if on any input and at each point of time it has at most ℓ open bins.

For some of the heuristics mentioned above there exist straightforward adaptions to the case of cardinality-constrained and class-constrained bin packing. For example Next-Fit packs the items in the currently open bin as long as it is possible, that is the total size of the items is at most 1 and there are at most k items or items of at most k different colors in the bin. Correspondingly, First-Fit packs the item into the first opened bin, which allows to accommodate this item.

Furthermore, for class-constrained bin packing there is another approach how to reuse an algorithm A, which is intended for classical bin packing: The ColorSets-approach partitions the set of colors into groups of size k. The first k distinct colors in the input form the first group, the second k distinct colors the second group, and so on. Then, the algorithm packs different groups separately using algorithm A. We denote this algorithm by $\text{CS}\,[A]$.

For classical and class-constrained bin covering there are three relevant heuristics:

- Dual-Next-Fit (DNF): DNF packs all arriving items into the same bin until the bin is covered. Then, the next items are packed into a new bin until this bin is covered, and so on.
- Dual-Harmonic DH_M: The algorithm DH_M is for the classical bin covering problem only. We assume without loss of generality that item sizes stem from the interval $[0, 1)$, otherwise we pack items of size one separately. Then, the interval $[0, 1)$ is partitioned into the subintervals $[0, 1/M), [1/M, 1/(M-1)), \dots, [1/2, 1)$. This partition induces also a partition of the set of items into M classes. DH_M packs items from different classes into different bins and it runs DNF independently for each class. That is, for $j \in \{2, \dots, M\}$ it uses exactly j items from the interval $[1/j, 1/(j-1))$ to cover a bin.
- FF2: This algorithm is for class-constrained bin covering in the case of unit sized items. That is we assume that each item – besides its color – has size 1 and the bin demand is given by an integer $B \in \mathbb{N}$. FF2 is based on a First-Fit approach: The algorithm adds each arriving item to the first bin for which it is suitable. To define the notion of suitable, consider a bin that contains already items with $k - t$ different colors. If this bin contains fewer than $B - t$ items, every item is suitable. Otherwise, if the number of items is exactly $B - t$, an item is only suitable if it has a color that is not yet contained in the bin.

1.3 Overview of Related Literature

Before we present our results in Section 1.4 we give an overview of the related literature. As bin packing is a fundamental problem in theoretical computer science there is a vast body of literature. Here, we mention only the results that are most important or closely related to the results obtained in this thesis. An extensive survey can be found in [19].

1.3.1 Classical Bin Packing

We begin with the classical bin packing problem in the offline case: Classical bin packing is a computationally hard problem, that is it is not only NP-complete [44] but also *strongly*

NP-complete [67]. Moreover, under the assumption $\mathsf{P} \neq \mathsf{NP}$ a reduction from the Partition-problem shows there cannot exist a polynomial-time algorithm with $\mathrm{AR}_a(A) < 3/2$, and the ratio $3/2$ is obtained by the algorithm First-Fit-Decreasing [79]. Considering the problem in an asymptotic sense it becomes easier: In the beginning of the 80s Fernandez de la Vega and Lueker proved the existence of an *asymptotic polynomial-time approximation scheme* [29], that is for every ϵ greater than zero there exists an algorithm A_ϵ with $\mathrm{AR}(A_\epsilon) \leq 1 + \epsilon$ and a polynomial runtime in the number of items, depending on $1/\epsilon$. Shortly afterwards Karmarkar and Karp showed that there is also an asymptotic *fully* polynomial-time approximation scheme [53]. That is, the runtime of A_ϵ is bounded polynomially in the number of items *and* $1/\epsilon$. It also follows from their approach that it is possible to construct an algorithm with polynomial runtime, which needs at most $\mathrm{OPT}(I) + \mathcal{O}\left(\log(\mathrm{OPT}(I))^2\right)$ many bins. Recently, there has been some progress in this direction: 2013 Rothvoss showed the existence of a randomized algorithm with polynomial expected runtime that needs at most $\mathrm{OPT}(I) + \mathcal{O}\left(\log(\mathrm{OPT}(I)) \log(\log(\mathrm{OPT}(I)))\right)$ many bins [73]. Shortly afterwards Hoberg and Rothvoss improved this result and proved the existence of a randomized algorithm with polynomial runtime, which needs at most $\mathrm{OPT}(I) + \mathcal{O}\left(\log(\mathrm{OPT}(I))\right)$ many bins [47].

We now turn to classical bin packing in the online setting. At first we consider the analysis in the worst case, that is investigations regarding the competitive ratio. We start giving results for some fundamental heuristics. A few of them were already mentioned in the previous section about algorithms.

- The algorithms Next-Fit and Worst-Fit were shown to be 2-competitive in the beginning of the 70s [50, 51].
- In the same decade it was found out that First-Fit and Best-Fit are 1.7-competitive [81]. Recently, it was shown that this also holds true for the absolute competitive ratio [31, 32].

Apart from these results for special natural-looking heuristics there has been a chase for the best online algorithm:

- The algorithm Refined-First-Fit proposed by Yao 1980 in [84] obtains a competitive ratio of $5/3$.
- Lee and Lee showed 1985 that the algorithm Refined-Harmonic, which is based on the Harmonic-approach, is approximately 1.6359-competitive [58].
- There were further improvements in [70] and [74] based on the same approach, which obtained competitive ratios of roughly 1.612 and 1.5889.
- Recently, Balogh et al. showed that their algorithm Advanced-Harmonic obtains a competitive ratio of 1.5783 [8].

Those upper bounds are complemented by a sequence of improved lower bounds for the competitive ratio:

- In the beginning of the 80s a lower bound of 1.5 was shown by Yao [84], which was independently improved to 1.5363 by Brown and Liang in [14] and [61].
- 1992 van Vliet proved a lower bound of 1.5401 [82].
- 2012 Balogh et al. refined the approach of van Vliet and gave an improved result of $248/161 \approx 1.5403$ in [10].
- Recently, Balogh et al. improved the bound to $(1363 - \sqrt{1387369})/120 \approx 1.5427$ [9].

Usually, these bounds carry over to randomized algorithms [16].

Apart from that Lee and Lee showed a lower bound of $h_\infty \approx 1.691$ for the competitive ratio of bounded-space algorithms, and also showed that $\mathrm{CR}(\mathrm{HARMONIC}_M)$ tends to this bound as M tends to infinity [58]. Interestingly, the algorithm Best-Fit is still 1.7-competitive if it is restricted to only two open bins [25].

Now we want give an overview of results in classical bin packing in the online setting, if the input is random. We start with results regarding the average performance ratio with respect to specific distributions. An important special case is that the items are distributed according to $\mathcal{U}[0,1]$, that is the uniform distribution on the interval $[0,1]$.

- For Next-Fit it holds $\text{APR}(\text{NF}, \mathcal{U}[0,1]) = 4/3$ [23].
- The algorithm Smart-Next-Fit has the same competitive ratio as Next-Fit, but if the items are drawn independently and identically distributed we perform better, that is we have $\text{APR}(\text{SNF}, \mathcal{U}[0,1]) \approx 1.227$ [69].
- Moreover, it was proved that $\text{APR}(\text{HARMONIC}_M, \mathcal{U}[0,1]) \approx 1.2899$ [57].

For the algorithms First-Fit and Best-Fit it can be shown that $\text{APR}(\text{FF}, \mathcal{U}[0,1]) = 1$ and $\text{APR}(\text{BF}, \mathcal{U}[0,1]) = 1$. To distinguish both algorithms better the expected waste of an algorithm A is introduced, that is $W_n^F(A) := \mathbb{E}\left[A(I_n^F) - S(I_n^F)\right]$, where $S(I_n^F)$ denotes the total sum of all items in I_n^F. For this quantity it was shown that $W_n^{\mathcal{U}[0,1]}(\text{FF}) \in \Theta(n^{2/3})$ [22, 78], and $W_n^{\mathcal{U}[0,1]}(\text{BF}) \in \Theta(\sqrt{n}\log(n)^{3/4})$ [59, 78]. We cannot improve substantially on these results: For any online algorithm A it holds that $W_n^{\mathcal{U}[0,1]}(A) \in \Omega\left(\sqrt{n\log(n)}\right)$ [78]. Moreover, for an M-bounded-space algorithm A we have $\text{APR}(A, \mathcal{U}[0,1]) \geq 1 + \frac{1}{4M+4}$ [25]. Additionally, there are more results concerning the analysis of *discrete uniform distributions* in [1, 18, 21, 55].

Now we turn to the case of probabilistic statements that do not depend on specific distributions. For the average performance ratio there are the following results:

- Rhee and Talagrand gave in 1993 a randomized online algorithm A with $\text{APR}(A, \mathcal{D}) = 1$, where $\mathcal{D}$ is the set of *all* distributions on $[0,1]$ [72]. This result is of theoretical interest, but since the algorithm has to compute frequently optimal solutions of subsets of the drawn items, it is not applicable in practice.
- In a setting, where the item sizes are integers and the bin capacity is given by an integer B, Csirik et al. presented the Sum-Of-Squares algorithm SS, which satisfies $\text{APR}(\text{SS}, \mathcal{P}) = 1$, where $\mathcal{P}$ denotes the set of all *perfect-packing distributions* [27]. This heuristic is a linear-time algorithm, and hence also of practical interest. Moreover, they provide an algorithm which is optimal for all distributions. But this algorithms exhibits again a large runtime.
- Later on, also in the setting of integer item sizes, Gupta and Radovanovic gave two optimal algorithms, which are based on Lagrangian relaxation [46].

With respect to the random-order ratio there are only two results published:

- Kenyon showed in 1996 that $1.08 \leq \text{RR}(\text{BF}) \leq 1.5$. So Best-Fit behaves better in expectation than in the worst case if the items arrive in random order.
- Moreover, 2008 it was proven that $\text{RR}(\text{NF}) = 2$ [20]. That is, Next-Fit does not benefit from such randomization compared to the worst case.

1.3.2 Cardinality-constrained and Class-constrained Bin Packing

As classical bin packing is an old fundamental problem, lots of variants were studied. In this thesis we consider cardinality-constrained and class-constrained bin packing. To the best of our knowledge these variants are up to now only studied with respect to the worst case.

We begin with considering cardinality-constrained bin packing: This problem was introduced 1975 in [56]. It is also NP-complete, but an APTAS and later on an AFPTAS were found [15, 39].

Babel et al. presented in 2004 two 2-competitive online algorithms [6]. Recently, Balogh et al. showed that there is no hope for better results, that is there exists a lower bound of 2 for the competitive ratio [7]. Apart from these results, the First-Fit heuristic is studied in [30] and the bounded-space setting in [36].

The class-constrained bin packing problem was introduced in [76]. It is NP-hard even in the case of unit sized items and $k = 2$, and strongly NP-hard if $k \geq 3$ [45, 77]. Several approximation and non-approximability results for the offline version are given in [38, 83].

The online version was studied at first in [77] for the case of unit sized items. Especially, there were presented two 2-competitive algorithms and a corresponding lower bound of 2. Furthermore, a lower bound of $5/3$ is given in the special case $k = 2$.

For arbitrary item sizes Epstein et al. gave a 2.635-competitive algorithm in [38] and Balogh et al. showed a lower bound of 1.717608 in the special case $k = 2$. Further results for specific heuristics were given in [38, 83].

1.3.3 Classical Bin Covering and Class-constrained Bin Covering

The classical bin covering problem is somewhat of the dual problem to classical bin packing. It was introduced in the beginning of the 80s in [4] and [5]. The problem is also NP-complete and the Dual-Next-Fit algorithm obtains an approximation ratio of $1/2$. Additionally, approximation algorithms with factors $2/3$ and $3/4$ were presented. 2001 an APTAS for the problem was found by Csirik et al. [26]. This result was improved a short time later by Jansen and Solis-Oba, who presented an AFPTAS [49].

The story of online bin covering is short: 1988 Csirik and Totik showed that no online algorithm for bin covering can obtain a competitive ratio better than $1/2$ [28]. As already the simple 1-bounded space algorithm DNF obtains this ratio, there is only very limited research in this area. In [17] the algorithm Dual-Harmonic is considered and it is also shown that *any* reasonable algorithm for online bin covering has a competitive ratio of $1/2$.

For random input it was shown that $\mathrm{APR}(\mathrm{DNF}, \mathcal{U}[0,1]) = 2/e \approx 0.7141$ [24]. Christ et al. compared Dual-Next-Fit and Dual-Harmonic for various performance measures, also probabilistic ones, and showed that $\mathrm{APR}(\mathrm{DH}_2, \mathcal{U}[0,1]) \approx 0.7141$, and $\mathrm{RR}(\mathrm{DNF}) \leq 4/5$ [17]. Finally, it was investigated how Markov chains can be used to design bounded-space algorithms for online bin covering [3].

Class-constrained bin covering is only studied in [37] and the authors restrict themselves to the case of unit sized items. For the offline version it was shown that there exists a simple optimal polynomial-time algorithm. For the online version an algorithm with competitive ratio of $\Omega(1/k)$, and a logarithmic upper bound of $\mathcal{O}(1/\log(k))$ are presented. Additionally, it is shown that $\mathrm{CR}(\mathrm{FF2}) = 1/B$.

1.4 Overview of Results, Outline and Bibliographical Notes

In Section 2 we give an overview of the most important stochastic techniques used in this thesis. Section 3, 4 and 5 contain our results with respect to packing and covering problems. The last section addresses open problems and possible further lines of research.

We have seen in the previous section that algorithms for packing and covering problems were analyzed on random input several times. But there are only few results where algorithms are analyzed with respect to the average performance ratio for larger classes of distributions or in the random-order model. In Section 3-5 we make progress in this direction.

We describe those results in more details:

- In Section 3 we investigate the complexity of classical, cardinality-constrained and class-constrained bin packing with respect to the introduced probabilistic performance measures. Rhee and Talagrand have shown in [72] the existence of an algorithm A for classical bin packing with $\mathrm{APR}(A, \mathcal{D}) = 1$, where $\mathcal{D}$ denotes the set of *all* probability distributions on $[0, 1]$. We adapt their approach to show the following results:
  - For cardinality-constrained bin packing there exists a randomized algorithm A with $\mathrm{RR}(A) = 1$, and therefore also $\mathrm{APR}(A, \mathcal{D}) = 1$.
  - For classical bin packing there exists for every ϵ greater than zero a randomized algorithm A_ϵ with $\mathrm{RR}(A_\epsilon) \leq 1 + \epsilon$. This is a step to the answer of the question raised in [54] about the existence of an asymptotic optimal algorithm for classical bin packing with respect to the random-order ratio.
  - Finally, for class-constrained bin packing there exists a randomized algorithm A with $\mathrm{APR}(A, \mathcal{D}) = 1$, where $\mathcal{D}$ denotes the class of all probability distributions on $[0, 1] \times \mathbb{N}$.

  These positive results are complemented by the following negative results obtained:
  - For classical bin packing we study another stochastic model, namely the *partial-permutations model*. Here, we are given a parameter $p \in [0, 1]$, which is a probability. Then, an opponent is allowed to design an instance I for classical bin packing. Afterwards, a subset of the items is chosen by flipping a coin with success probability p for each item. Then, all chosen items are randomly permuted. For p equal to one we obtain the random-order model, and for p equal to zero the opponent is allowed to give a worst-case sequence. Hence, the partial-permutation model interpolates between both cases.
    We show that for every $p \in [0, 1)$ there exists an ϵ_p greater than zero such that for every algorithm A there exist arbitrary large instances I such that the expected number of opened bins by A on I is at least $(1 + \epsilon_p)\,\mathrm{OPT}(I)$. So, if the adversary is able to control a constant fraction of the input, it is not possible to obtain a positive result as in the case where p is equal to one.
  - Furthermore, we show that $\mathrm{RR}(A) \geq 10/9$ for all online algorithms A for class-constrained bin packing. This demonstrates that the complexity of class-constrained bin packing differs with respect to the average performance ratio and the random-order ratio.
- In Section 4 we deal with bounded-space algorithms for classical bin packing and classical bin covering. For classical bin packing Lee and Lee showed in [58] that no bounded-space algorithm – independent of the allowed number of open bins – can obtain a competitive ratio better than $h_\infty \approx 1.691$. We show that if the items arrive in random order it is possible to break this bound. That is, we construct an algorithm, which is based on the Harmonic-algorithm, that uses only four open bins and obtains a random-order ratio of at most 1.671. Moreover, in the past it was already shown that the random-order ratio of the 1-bounded-space algorithm Next-Fit is 2. We generalize this result to more algorithms, and show that it is also true for the average performance ratio.
  For classical bin covering we show that the algorithm Dual-Next-Fit has a random-order ratio of at least 0.502. This is in contrast to the worst-case, as it was shown that no online algorithm can obtain a competitive ratio better than 0.5.
- We conclude the probabilistic analysis of packing and covering problems with analyzing selected established heuristics for class-constrained bin packing and bin covering.
  In the analysis of class-constrained bin packing we observe again that the results for the average performance ratio and random-order ratio differ for many studied algorithms. So,

the different asymptotic complexity of this problem, already observed in Section 3, also plays a role for established heuristics. Especially, we consider the ColorSets-approach: Here, we show that the average performance ratio of CS $[A]$ cannot be better than h_∞, independently of the chosen algorithm A. Furthermore, this bound is almost tight, that is we can find a family of algorithms whose average performance ratios tends to this bound. However, in the random-order model we obtain a lower bound of 2.
Considering class-constrained bin covering we do not observe a difference between the two probabilistic performance measures. Among other results, we show that $\mathrm{RR}(\mathrm{FF2}) = 1$ in the case of unit sized items. This is in contrast to a competitive ratio of $1/B$. Then, we use this result to establish an online algorithm with random-order ratio of $1/3$ for class-constrained bin covering with general item sizes. It follows that there exists also a randomized offline algorithm with approximation ratio of $1/3$. To our knowledge this is the first offline algorithm for this problem in the case of arbitrary item sizes presented.

Preliminary versions of the results concerning the analysis of the Dual-Next-Fit algorithm in Section 4 and the analysis of selected heuristics for class-constrained bin packing and bin covering in Section 5 have been published at conferences and as a technical report:

- Carsten Fischer and Heiko Röglin. Probabilistic analysis of the dual next-fit algorithm for bin covering. *arXiv preprint arXiv:1512.04719*, 2015
- Carsten Fischer and Heiko Röglin. Probabilistic analysis of the dual next-fit algorithm for bin covering. In *Proc. of the Latin American Symposium on Theoretical Informatics (LATIN)*, pages 469–482, 2016.
- Carsten Fischer and Heiko Röglin. Probabilistic analysis of online (class-constrained) bin packing and bin covering. In *Proc. of the Latin American Symposium on Theoretical Informatics (LATIN)*, pages 461–474, 2018.

2 Stochastic Background

Most of the results presented in Section 3, 4 and 5 are probabilistic statements. To obtain these results we will mainly use tools from the fields *concentration inequalities*, *Markov chains*, and *stochastic upright matchings*.

Usually, if we deal with random instances we hope that they satisfy with high probability a certain structure, and that we can exploit this structure to make statements about the behavior of the algorithm. Concentration inequalities are an important tool to quantify bounds for the probability of certain realizations of our random instances.

Another important field is the area of Markov chains. A Markov chain – in our context – is a stochastic process $(X_n)_{n\in\mathbb{N}_0}$, where the random variables X_n live on a state space $\mathcal{S}$. Here, the outcome of X_{n+1} depends only on X_n. Markov chains can be used to model for example the configurations an algorithm runs through on random input. Using the mathematical machinery for Markov chains this allows us to obtain statements about the long-term behavior of the analyzed algorithm. This procedure is a common approach and was used for example in [3] and [54].

The third important tool is the stochastic upright matching problem. In the stochastic upright matching problem several plus- and minus-points are randomly placed in the plane. Then, we are interested in how many plus-points we can match to minus-points such that the following two constraints are satisfied:

1. A plus-point p^+ can only be matched to a minus-point p^- that is located on the upper and right-hand side of p^+;
2. No two different plus-points can be matched to the same minus-point.

After the introduction of the problem it was discovered that it can be used in the analysis of bin packing algorithms on random input. This was used for example in [54] and [72].

Finally, in the last part of this section we will give two important identities for the probabilistic performance measures.

2.1 Concentration Inequalities

If we deal with worst-case analysis, that is analysis with respect to the competitive ratio the adversary is very powerful: He has full control over the items and their order of arrival. Consider the following input for the bin covering problem $I = (1-\epsilon, \dots, 1-\epsilon, \epsilon, \dots, \epsilon)$ consisting of $2n$ items of size $1-\epsilon$ and $2n$ items of size ϵ and apply the algorithm Dual Next-Fit. Here we choose ϵ such that $2n\epsilon < 1$. On this input it is clear that $\mathrm{OPT}(I) = 2n$ and $\mathrm{DNF}(I) = n$, so Dual Next-Fit cannot be better than $1/2$-competitive with respect to the competitive ratio. But if we take a close look at I we see that this sequence is somehow artificial: If we call items of size $1-\epsilon$ large, and items of size ϵ small, then our input starts with $2n$ large items and ends with $2n$ small items. When we consider probabilistic inputs we have the hope that artificial bad instances are unlikely to occur.

In our context if we talk about concentration inequalities we are usually interested in bounding the probability of deviating from the mean. So if Z is a real-valued random variable, we want to give bounds for the term

$$\mathbb{P}\left[|Z - \mathbb{E}\left[Z\right]| \geq \lambda\right], \tag{1}$$

where $\lambda > 0$.

One of the most fundamental inequalities is *Markov's inequality*: Markov's inequality states that for a non-negative random variable X and $\lambda > 0$ it holds $\mathbb{P}[X \geq \lambda] \leq \frac{1}{\lambda} \cdot \mathbb{E}[X]$. To obtain a concentration inequality in the style of (1) we set $X := |Z - \mathbb{E}[Z]|$.

Markov's inequality can be used to obtain a stronger concentration bound, namely *Chebyshev's inequality*: Chebyshev's inequality states that for a random variable Z with $0 < \mathrm{Var}[Z] < \infty$ we have $\mathbb{P}[|Z - \mathbb{E}[Z]| \geq \lambda] \leq \frac{1}{\lambda^2} \mathrm{Var}[Z]$.

After this short introduction we study the special case when Z depends in some way on several independent random variables $X_1, \ldots, X_n$.

2.1.1 Independent and Identically Distributed Random Variables

In this thesis, especially when we consider the average performance ratio we often consider the case when Z depends in some way on several independent real-valued random variables $X_1, \ldots, X_n$, that is there exists a function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $Z = f(X_1, \ldots, X_n)$. This setup is studied extensively in the book [13].

An important case is that Z is the sum of the random variables $X_1, \ldots, X_n$, so we have $Z = \sum_{i=1}^n X_i$. Applying Chebyshev's inequality in this setting and using the independence assumption, we obtain

$$\mathbb{P}\left[\left|\sum_{i=1}^n X_i - \sum_{i=1}^n \mathbb{E}[X_i]\right| \geq \lambda\right] \leq \frac{1}{\lambda^2} \sum_{i=1}^n \mathrm{Var}[X_i]\,. \tag{2}$$

It turns out that in this setting there exist much stronger statements. One of the most famous is *Hoeffding's inequality* published by Wassily Hoeffding 1963 in [48]:

▶ **Proposition 5** (Hoeffding's Inequality). *Let $X_1, \ldots, X_n$ be real-valued independent random variables such that X_i takes its values in $[a_i, b_i]$ almost surely for all $i \in [n]$, where $a_i, b_i \in \mathbb{R}$. Then, for every $\lambda > 0$, we have*

$$\mathbb{P}\left[\sum_{i=1}^n X_i \geq \sum_{i=1}^n \mathbb{E}[X_i] + \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

and

$$\mathbb{P}\left[\sum_{i=1}^n X_i \leq \sum_{i=1}^n \mathbb{E}[X_i] - \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

In the following example we compare the bounds given by Chebyshev's and Hoeffding's inequality:

▶ **Example 6.** We think back of the instance, presented as worst-case instance for DNF, $I = (1-\epsilon, \ldots, 1-\epsilon, \epsilon, \ldots, \epsilon)$ consisting of $2n$ large items followed by $2n$ small items. As the number of large and small items is equal a reasonable stochastic setting could be the following: We draw independently $4n$ items, where the i-th item is large with probability $1/2$, and small otherwise. Let N denote the random variable that represents the number of drawn large items among the first $2n$ items. If we set X_i equal to one if the i-th drawn item is large, and zero otherwise, we have $N = \sum_{i=1}^{2n} X_i$. While in I the number of large items among the first $2n$ drawn items is equal to $2n$ the expected number $\mathbb{E}[N]$ of large items among the first $2n$ drawn items is equal to n. But what is the probability that an extreme event occurs, for example that N deviates from $\mathbb{E}[N]$ by at least the half of the expected value?

We have $\mathrm{Var}[X_i] = 1/4$ and $\lambda^2 = (\mathbb{E}[N]/2)^2 = n^2/4$. Therefore, using (2), we obtain

$$\mathbb{P}\left[|N - \mathbb{E}[N]| \geq \frac{1}{2} \cdot \mathbb{E}[N]\right] \leq \frac{2}{n}.$$

On the other hand, we can apply Hoeffding's inequality as we have $0 \leq X_i \leq 1$ for all i. Then, we have

$$\mathbb{P}\left[|N - \mathbb{E}[N]| \geq \frac{1}{2} \cdot \mathbb{E}[N]\right] \leq 2\exp\left(-\frac{1}{2} \cdot n\right).$$

As we see the bound obtained by applying Hoeffding's inequality is for large n much sharper and tends to zero exponentially.

Sometimes we will encounter the case that the sum of the random variables will be much smaller than the number of random variables that we add up. Then, *Bernstein's inequality*, which depends on the second moments of the X_i, will be helpful (see e.g. Section 2.7 and 2.8 in [13]):

▶ **Proposition 7** (Bernstein's Inequality). *Let $X_1, \dots, X_n$ be real-valued independent random variables such that X_i is upper bounded by M for all $i \in [n]$. Then, for every $\lambda > 0$, we have*

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}[X_i] + \lambda\right] \leq \exp\left(-\frac{\lambda^2}{2\left[\sum_{i=1}^{n} \mathbb{E}[X_i^2] + M\lambda/3\right]}\right).$$

If $X_1, \dots, X_n$ are real-valued independent random variables such that X_i is bounded from below by $-M$ for all $i \in [n]$, then for every $\lambda > 0$ it holds

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \leq \sum_{i=1}^{n} \mathbb{E}[X_i] - \lambda\right] \leq \exp\left(-\frac{\lambda^2}{2\left[\sum_{i=1}^{n} \mathbb{E}[X_i^2] + M\lambda/3\right]}\right).$$

▶ **Example 8.** We take up the previous example. Let us now assume that the number n of drawn items is at least 100, and we draw a large item with probability $100/n$. Then, the expected number of drawn large items $\mathbb{E}[N]$ is equal to 100. At first we calculate the probability deviating from this value by at least 50 using Hoeffding's inequality. We obtain

$$\mathbb{P}\left[|N - \mathbb{E}[N]| \geq \frac{1}{2}\mathbb{E}[N]\right] \leq 2\exp\left(-\frac{1}{n} \cdot 5000\right).$$

We see that this estimate tends to one as n tends to infinity, and therefore is not helpful.

Now we apply Bernstein's inequality: The random variables X_i are still bounded from above by 1 and for the second moment we have $\mathbb{E}\left[X_i^2\right] = 100/n$. Then, we obtain

$$\mathbb{P}\left[|N - \mathbb{E}[N]| \geq \frac{1}{2}\mathbb{E}[N]\right] \leq 2\exp\left(-\frac{75}{7}\right) \approx 0.0004.$$

Hoeffding's and Bernstein's inequality help us to control the deviation for the sum of n random variables. But in some cases we need to control the deviations of *all* partial sums simultaneously. The following maximal inequality, proved e.g. in [66], allows us to reduce this question to bounding the deviation of the entire sum.

▶ **Proposition 9** (Maximal Inequality). *Let $X_1, \dots, X_n$ be independent and identically distributed real-valued random variables. Then, we have*

$$\mathbb{P}\left[\max_{1 \leq j \leq n}\left|\sum_{i=1}^{j} X_i\right| > \lambda\right] \leq C \cdot \mathbb{P}\left[\left|\sum_{i=1}^{n} X_i\right| > \lambda/C\right],$$

whenever $\lambda \geq 0$, where C is a universal constant.

Now, we turn towards the more general case that $Z = f(X_1, \ldots, X_n)$: If no argument of f is of overwhelming importance, then Z satisfies also nice concentration properties. We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ has the *bounded difference property* if there exist nonnegative constants $c_1, \ldots, c_n$ such that for all $i \in [n]$

$$\sup_{x_1,\ldots,x_n,x_i'} |f(x_1, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)| \leq c_i.$$

Then, the *Bounded Difference Inequality* also called *McDiarmid's inequality* holds true:

▶ **Proposition 10** (Bounded Difference Inequality). *Let $X_1, \ldots, X_n$ be real-valued independent random variables, and $f : \mathbb{R}^n \to \mathbb{R}$ be a function, which satisfies the bounded difference property with constants $c_1, \ldots, c_n$. Then,*

$$\mathbb{P}\left[f(X_1, \ldots, X_n) \geq \mathbb{E}\left[f(X_1, \ldots, X_n)\right] + \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^n c_i^2}\right)$$

and

$$\mathbb{P}\left[f(X_1, \ldots, X_n) \leq \mathbb{E}\left[f(X_1, \ldots, X_n)\right] - \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^n c_i^2}\right).$$

The bounded difference inequality shows for example that for an arbitrary distribution F on $[0, 1]$ and items $X_1, \ldots, X_n$ drawn independently according to F, the value of $\mathrm{OPT}((X_1, \ldots, X_n))$ is highly concentrated around its mean.

Finally, especially when we design (nearly) optimal algorithms in Section 3, we will also need a concentration statement for a different setup. Let F be an arbitrary distribution on $[0, 1]$ with cumulative distribution function $G\colon [0, 1] \to [0, 1]$. Moreover, let $X_1, \ldots, X_n$ be independent random variables that are distributed according to F. Then, we can approximate F by the empirical distribution $\hat{F}$ with empirical distribution function $\hat{G}_n(x) := \frac{1}{n}\sum_{i=1}^n I_{\{X_i \leq x\}}$. Here, $I_{\{X_i \leq x\}}$ denotes the indicator function, that means this function is equal to 1 if $X_i \leq x$ and 0 otherwise.

The *Dvoretzky-Kiefer-Wolfowitz Inequality* (in short: DKW inequality) shows that the difference between G and $\hat{G}_n$ vanishes exponentially fast [33, 64].

▶ **Proposition 11** (Dvoretzky-Kiefer-Wolfowitz Inequality). *Let F denote an arbitrary distribution on $[0, 1]$, $X_1, \ldots, X_n$ independent random variables distributed according to F, and let $\hat{G}_n$ denote the corresponding empirical distribution function. Then, we have for $\lambda > 0$ arbitrary*

$$\mathbb{P}\left[\sup_{x \in [0,1]} \left|\hat{G}_n(x) - G(x)\right| \geq \lambda\right] \leq 2\exp\left(-2n\lambda^2\right).$$

2.1.2 Dependent and Identically Distributed Random Variables

In the previous part we assumed that the random variables X_i are independent. When we analyze algorithms with respect to the random-order ratio we cannot make this assumption. In this case we have a probabilistic setup which is similar to drawing several items from an urn without replacing them. Fortunately, also in this setting several important concentration inequalities hold true.

That Hoeffding's inequality also holds true under these circumstances was already mentioned by Hoeffding himself in [48].

▶ **Proposition 12** (Hoeffding's Inequality for Sampling Without Replacement). *Let $\mathcal{X} = \{x_1, \dots, x_N\}$ be a finite population of N reals ($\mathcal{X}$ can be a multiset) and $X_1, \dots, X_n$ be a random sample drawn without replacement from $\mathcal{X}$. Let $a := \min_{1 \le i \le N} x_i$ and $b := \max_{1 \le i \le N} x_i$. Then, for all $\lambda > 0$,*

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] + \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{n(b-a)^2}\right)$$

and

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \leq \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] - \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{n(b-a)^2}\right).$$

Also Bernstein's inequality can be transposed to the new setting [12]:

▶ **Proposition 13** (Bernstein's Inequality for Sampling Without Replacement). *Let $\mathcal{X} = \{x_1, \dots, x_N\}$ be a finite population of N reals ($\mathcal{X}$ can be a multiset) and $X_1, \dots, X_n$ be a random sample drawn without replacement from $\mathcal{X}$. Let $a := \min_{1 \le i \le N} x_i$ and $b := \max_{1 \le i \le N} x_i$. Then, for all $\lambda > 0$,*

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] + \lambda\right] \leq \exp\left(-\frac{\lambda^2}{2\left[\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + (b-a)\lambda/3\right]}\right)$$

and

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \leq \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] - \lambda\right] \leq \exp\left(-\frac{\lambda^2}{2\left[\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + (b-a)\lambda/3\right]}\right).$$

In fact, in the setting of sampling without replacement we often have even stronger concentration properties than in the case of sampling with replacement. More sophisticated statements can be found for example in [12].

Finally, it is also possible to obtain a maximal inequality in the case of drawing items according to sampling without replacement as shown in [68].

▶ **Proposition 14** (Maximal Inequality for Sampling Without Replacement). *Let $X_1, \dots, X_n$ be an exchangeable sequence of real-valued random variables with $n \geq 2$. Then there exists an universal constant C such that*

$$\mathbb{P}\left[\max_{1 \le j \le n}\left|\sum_{i=1}^{j} X_i\right| > \lambda\right] \leq C \cdot \mathbb{P}\left[\left|\sum_{i=1}^{\lfloor n/2 \rfloor} X_i\right| > \lambda/C\right].$$

2.2 Markov Chains

In this part we will describe the most important statements in the area of Markov chains, which are important for this thesis. In order to do this will follow mostly the book [75]. Another nice overview – in the case of finite state space Markov chains – is given in [60].

2.2.1 Basics

A *time-discrete stochastic process* $(X_n)_{n \in \mathbb{N}_0}$ is a vector of random variables that attain values in a *state space* $\mathcal{S}$ and are realized on the same probability space. We say that X_n is the

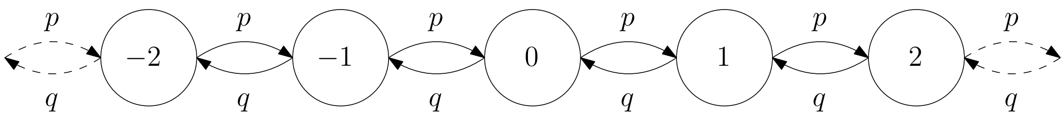

Figure 1 This is the transition graph of a random walk on $\mathbb{Z}$, where we go to the right with probability p and to the left with probability $q := 1 - p$. If $p = q$, this is a symmetric random walk.

state of the process at time n. In our applications we will usually deal with the case that $\mathcal{S}$ is a countable or finite set.

We say that a stochastic process $(X_n)_{n\in\mathbb{N}_0}$ is a *time-homogeneous Markov chain* if we have for arbitrary $n \in \mathbb{N}_0$ and $s \in \mathcal{S}$

$$\mathbb{P}\left[X_{n+1} = s \mid X_1, \ldots, X_n\right] = \mathbb{P}\left[X_{n+1} = s \mid X_n\right].$$

This is the so-called *Markov property* and means that the outcome of the next state X_{n+1} depends only on the state X_n, and is independent of any other states attained in the past.

We assume that $\mathcal{S} = \{s_1, s_2, \ldots, s_m\}$ for an $m \in \mathbb{N}$ in the case that $\mathcal{S}$ is finite, and $\mathcal{S} = \{s_i : i \in \mathbb{N}\}$ in case that $\mathcal{S}$ is countably infinite.

For $s, \tilde{s} \in \mathcal{S}$ we denote by $p_{s,\tilde{s}}$ the probability to go from state s to $\tilde{s}$ in one step. Then $P = (p_{s,\tilde{s}})_{s,\tilde{s}\in\mathcal{S}}$ denotes the *transition matrix* and for an arbitrary s we have $\sum_{\tilde{s}\in\mathcal{S}} p_{s,\tilde{s}} = 1$. So, if $\mathcal{S}$ is countable our transition matrix is an infinite matrix. It is often helpful to visualize the transition matrix as a *transition graph*. Then, the set of nodes is given by $\mathcal{S}$ and between each two states s and $\tilde{s}$ with $p_{s,\tilde{s}} > 0$ there is an edge with weight $p_{s,\tilde{s}}$.

Furthermore, the starting state X_0 of the Markov chain is determined by a probability distribution μ on $\mathcal{S}$. In case that we have $\mu = \delta_s$ for an $s \in \mathcal{S}$ (where δ_s denotes the dirac measure) we say that the Markov chain starts in state s.

▶ **Example 15.** We present several Markov chains to give an idea of this concept:

1. An important class of Markov chains are sums of independent, identically distributed random variables. For example, let $Y_1, Y_2, \ldots$ be independent random variables with $\mathbb{P}\left[Y_i = 1\right] = p$ and $\mathbb{P}\left[Y_i = -1\right] = 1 - p$, and we set $X_n := \sum_{i=1}^{n} Y_i$. The transition graph of this Markov chain is described in Figure 1. Sometimes we will call Markov chains of this type *random walks*, but we will also call similar Markov chains random walks. In case that $p = q$, the expected value of the random variables Y_i is equal to zero. In this case we will call the random walk *symmetric*.
2. It is possible to extend the concept of the symmetric one-dimensional random walk to higher dimensions, that is to the state space $\mathbb{Z}^d$. Then, the increments are drawn with respect to the uniform distribution from the set of the canonical unit vectors and their negative counterparts. The transition graph in case $d = 2$ is given in Figure 2.
3. Finally, we present another Markov chain, where the transition graph is shown in Figure 3. Here, the state space is $\mathbb{N}_0$ and we have a drift towards zero.

2.2.2 Properties of Markov Chains

In this part we introduce some important properties of Markov chains. We begin with the notion of irreducibility:

▶ **Definition 16.** A Markov chain is *irreducible* if for two arbitrary states $s, \tilde{s} \in \mathcal{S}$ there exists a $u \in \mathbb{N}_0$ such that it is possible to go from s to $\tilde{s}$ in u steps with a positive probability.

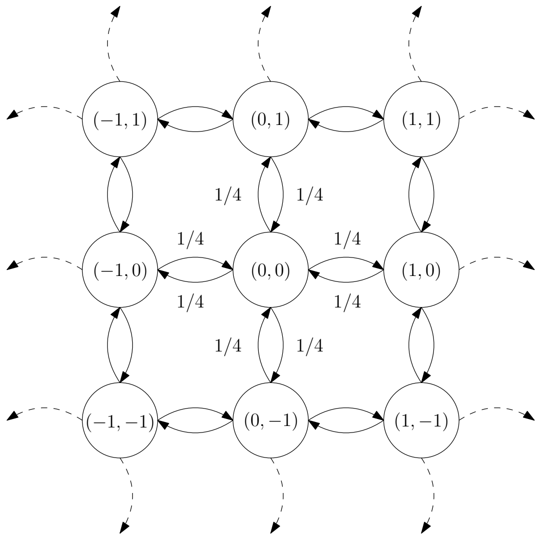

Figure 2 This is the transition graph of a symmetric random walk on $\mathbb{Z}^2$.

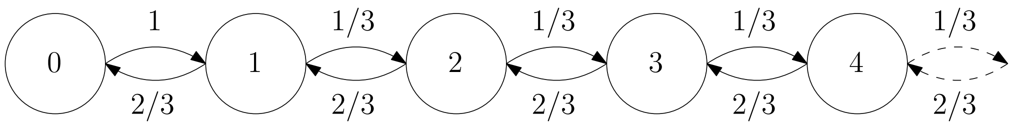

Figure 3 This is the transition graph of a Markov chain on $\mathbb{N}_0$, where we have a drift towards zero.

Moreover, in the following we use two abbreviations: Let $\mathbb{P}_s(A) := \mathbb{P}[A \mid X_0 = s]$ and $\mathbb{E}_s[Z] := \mathbb{E}[Z| X_0 = s]$.

The *hitting time* T_s is defined as $T_s := \min\{n \geq 1 : X_n = s\}$. In case that our Markov chain starts in state s we refer to this also as *first return time*.

Based on the notion of the first return time we define several properties of states:

▶ **Definition 17.** We say that state s is *transient* if we have $\mathbb{P}_s[T_s < \infty] < 1$, that is if the Markov chain starting in s returns to s with probability smaller than one. Otherwise, we call the state *recurrent*. Moreover, we say that a recurrent state s is *positive recurrent* if $\mathbb{E}_s[T_s] < \infty$, and *null recurrent* if $\mathbb{E}_s[T_s] = \infty$.

Interestingly, if the Markov chain is irreducible then all states behave in the same way:

▶ **Proposition 18.** *If we have an irreducible Markov chain, then all states $s \in \mathcal{S}$ share the same behavior, that is they are either transitive, positive recurrent or null recurrent.*

▶ **Example 19.** We take up again the Markov chains considered in Example 15. Looking at the transition graphs it is easy to see that all of the three Markov chains are irreducible. At first we consider the random walk on $\mathbb{Z}$, where we go to the right with probability p and to the left with probability $1 - p$. In case that we have a symmetric random walk, that is we have $p = 1/2$, this random walk is null recurrent. If $p \neq 1/2$, then the Markov chain has a drift to the left or to the right. In this case the random walk is transient. More generally it can be shown that the symmetric random walk on $\mathbb{Z}^d$ is null recurrent if $d \leq 2$, but transient if $d \geq 3$.

Finally, we consider the Markov chain with transition graph depicted in Figure 3. For this Markov chain it is possible to show that it is positive recurrent.

2.2.3 Stationary Distributions and Long-term Averages

Since we usually deal with asymptotic performance measures, we are mostly interested in the long-term behavior of Markov chains. Of special interest is the question which conditions a function $f\colon \mathcal{S} \to \mathbb{R}$ and a Markov chain (X_n) with transition matrix P have to satisfy such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[f(X_i)\right]$$

exists, and in case of existence how we can compute this limit.

An important concept will be stationary distributions.

► **Definition 20.** A probability measure π on $\mathcal{S}$ is a *stationary distribution* for the Markov chain (X_n) with transition matrix P if we have for all $s \in \mathcal{S}$

$$\pi(s) = \sum_{\tilde{s}\in\mathcal{S}} \pi(\tilde{s}) p_{\tilde{s},s}.$$

It follows from this definition that if X_0 is distributed according to a stationary distribution π, then X_i is distributed according to π for all $i \in \mathbb{N}$. Furthermore, it follows from this definition that we can determine a stationary distribution π by calculating the left-eigenvector to eigenvalue 1 of the transition matrix P.

While Markov chains with finite state spaces always possess at least one stationary distribution, this is not true in the case that $\mathcal{S}$ is infinite. The following statement gives conditions for the existence of stationary distributions and shows an important connection of the stationary distribution and the first return time.

► **Proposition 21.** *An irreducible Markov chain (X_n) possesses a stationary distribution if and only if all of its states are positive recurrent. In that case, the stationary distribution is unique, and independent of the distribution of X_0.*

Let π denote this distribution, and let $\pi(s)$ denote the probability of being in state s. Then, we have $\pi(s) = 1/\mathbb{E}_s\left[T_s\right]$.

► **Example 22.** We take up again the Markov chains presented in our two previous examples. It follows from Proposition 21 that only the Markov chain with drift towards zero possesses a stationary measure. Now, we will compute the stationary distribution π for this chain. In order to do this, we will solve an infinite system of linear equations, which is based on Definition 20. We have

$$\begin{aligned}
\pi(0) &= \frac{2}{3}\pi(1) \\
\pi(1) &= \pi(0) + \frac{2}{3}\pi(2) \\
\pi(2) &= \frac{1}{3}\pi(1) + \frac{2}{3}\pi(3) \\
\pi(3) &= \frac{1}{3}\pi(2) + \frac{2}{3}\pi(4) \\
&\dots
\end{aligned}$$

Solving this system and normalizing the solution we obtain $\pi(0) = 1/4$, $\pi(1) = 3/8$, $\pi(2) = 3/16$, $\pi(3) = 3/32$, and so on.

Furthermore, under mild conditions, it is possible to compute long-term averages using an (existing) stationary distribution.

▶ **Proposition 23.** *Let X_n be an irreducible and positive recurrent Markov chain. Let π denote its unique stationary distribution. Furthermore, let $f\colon \mathcal{S} \to \mathbb{R}$ be a bounded function. Then, we have*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[f(X_i)\right] = \sum_{s\in\mathcal{S}} \pi(s) f(s).$$

We want to finish this subsection by presenting the procedure to obtain a lower bound for the random-order ratio of Best-Fit given in [54]. This approach illustrates the use of Markov chains in the probabilistic analysis of bin packing and bin covering problems in a nice way.

▶ **Example 24.** A possible way to obtain a lower bound for the random-order ratio of an algorithm A (in the context of bin packing) is to choose a concrete distribution F and to take advantage of the fact that

$$\mathrm{RR}(A) \geq \lim_{n\to\infty} \frac{\mathbb{E}\left[A(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}.$$

We will prove this relationship in Section 2.4.

Let F be the distribution on the set of items $\mathcal{I} = \{1/3, 1/2\}$, where we choose an item of size $1/2$ with probability p and an item of size $1/3$ with probability $1-p$. As Best-Fit is optimal if p is equal to zero or one we assume that $p \in (0,1)$.

We say that a bin is *open*, if it is possible for the algorithm to pack items arriving in the future in this bin. If Best-Fit operates on items with sizes in $\mathcal{I}$ we see that we can describe the current configuration of the algorithm with five states:

- there is no open bin (state s_a);
- there is one open bin containing an item of size $1/3$ (state s_b);
- there is one open bin containing an item of size $1/2$ (state s_c);
- there is one open bin containing two items of size $1/3$ (state s_d);
- there are two open bins; one containing two items of size $1/3$ and one containing one item of size $1/2$ (state s_e).

So let $\mathcal{S} = \{s_a, s_b, s_c, s_d, s_e\}$ denote the state space of our Markov chain, and the transition probabilities between the states are described in Figure 4.

In the following we want to calculate or to estimate the terms $\lim_{n\to\infty} \mathbb{E}\left[\mathrm{BF}(I_n^F)\right]/n$ and $\lim_{n\to\infty} \mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]/n$.

We start with the analysis of Best-Fit: It holds

$$\begin{aligned}\mathbb{E}\left[\mathrm{BF}(I_n^F)\right] &= \mathbb{E}\left[\sum_{i=1}^{n} I_{\{\text{BF opens a new bin for the } i\text{-th item}\}}\right]\\ &= \sum_{i=1}^{n} \mathbb{P}\left[\text{BF opens a new bin for the } i\text{-th item}\right].\end{aligned}$$

So let $f\colon \mathcal{S} \to [0,1]$ with

$$f(s) = \mathbb{P}\left[\text{BF opens a new bin for the next item being in state } s\right],$$

that is we have $f(s_a) = 1$, $f(s_b) = 0$, $f(s_c) = 0$, $f(s_d) = p$, and $f(s_e) = 0$.

As we can see in Figure 4 our Markov chain is irreducible, and since we have a finite state space it is also positive recurrent. Hence, it follows from Proposition 21 that there exists a unique stationary probability distribution $\pi\colon \mathcal{S} \to [0,1]$.

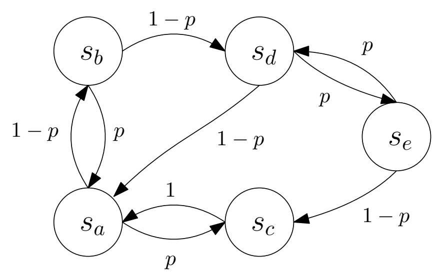

Figure 4 The transition graph of the Markov chain that describes the behavior of Best-Fit in Example 24.

Calculating the left-eigenvector to eigenvalue one of the transition matrix and normalizing, we obtain

$$\pi(s_a) = \frac{1+p}{3+3p-3p^2+p^3} \quad \text{and} \quad \pi(s_d) = \frac{1-p}{3+3p-3p^2+p^3}.$$

So, since f is a bounded function, we can apply Proposition 23 and obtain

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\mathrm{BF}(I_n^F)\right] = \frac{1+p}{3+3p-3p^2+p^3} \cdot 1 + \frac{1-p}{3+3p-3p^2+p^3} \cdot p$$
$$= \frac{1+2p-p^2}{3+3p-3p^2+p^3}.$$

Now, we want to estimate $\lim_{n\to\infty} \mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]/n$. Let N_n denote the number of items of size $1/2$ in I_n^F and M_n the number of items of size $1/3$. Then, we have

$$\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] \le \mathbb{E}\left[\left\lceil \frac{1}{2}N_n \right\rceil + \left\lceil \frac{1}{3}M_n \right\rceil\right] \le \mathbb{E}\left[\frac{1}{2}N_n + \frac{1}{3}M_n\right] + 2$$
$$= \left(\frac{1}{2}p + \frac{1}{3}(1-p)\right) n + 2 = \left(\frac{1}{3} + \frac{1}{6}p\right) n + 2.$$

So the limit can be bounded from above by $1/3 + p/6$.

Combining the result obtained for both limits, it follows that

$$\mathrm{RR}(\mathrm{BF}) \ge \frac{1+2p-p^2}{3+3p-3p^2+p^3} \cdot \frac{6}{2+p}.$$

Choosing $p \approx 0.456$ we obtain a lower bound for the random-order ratio which is slightly larger than 1.08.

2.3 Stochastic Upright Matchings

When we design and analyze algorithms that are (almost) optimal for stochastic performance ratios in Section 3 we will deal with problems called *stochastic upright matching problems* in literature. For our purpose it is more convenient to introduce these problems as stochastic *upleft* matching problems. An instance $\mathcal{P} = (\mathcal{P}^+, \mathcal{P}^-)$ for an upleft matching problem consists of two finite point sets $\mathcal{P}^+$ and $\mathcal{P}^-$ in $\mathbb{R}^2$, that is we have a set of points that are labeled with a *plus* and a set of points that are labeled with a *minus*. The aim is to match as many points from $\mathcal{P}^+$ to $\mathcal{P}^-$ as possible satisfying the following constraints:

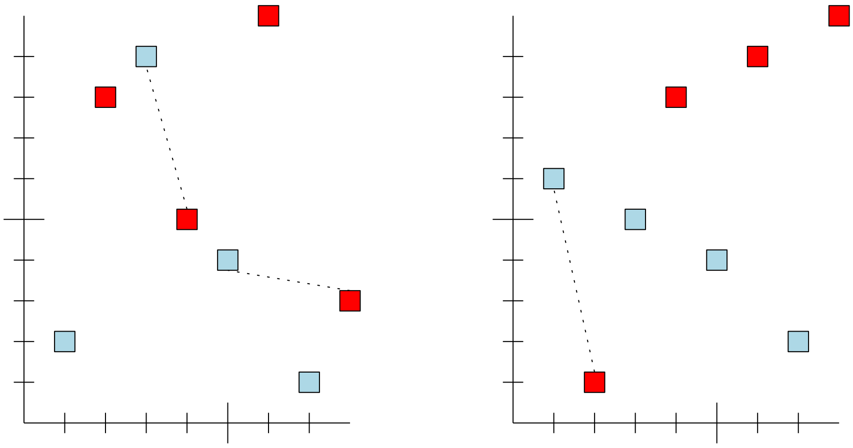

Figure 5 Here, two possible realizations of an instance for Matching-Variant 1 with parameters $k = 4$ and $n = 10$ are illustrated. The red squares represent the points from $\mathcal{P}^+$, and the blue ones the points from $\mathcal{P}^-$. The dashed lines indicate a maximum matching.

- A point $(x^+, y^+) \in \mathcal{P}^+$ can only be matched to a point $(x^-, y^-) \in \mathcal{P}^-$ with $x^- \leq x^+$ and $y^- \geq y^+$;
- It is not allowed to match two different points $p_1^+, p_2^+ \in \mathcal{P}^+$ to the same point p^- in $\mathcal{P}^-$.

We call a matching for $\mathcal{P}$ *maximum*, if it maximizes the number of matched points in $\mathcal{P}^+$ and we let $U(\mathcal{P})$ denote the number of unmatched points in $\mathcal{P}^+$ in a maximum matching. An example for an upleft matching instance is given in Figure 5.

We notice that the value of a maximum matching only depends on the *relative* positions of the points, and not on their absolute values. Therefore, in the following we assume without loss of generality that the x-coordinates are in $\{1, 2, 3, \ldots\}$.

These problems will be used in the following way in the analysis of (nearly) optimal algorithms: Let $\{t_1, \ldots, t_m\}$ denote the time points, when items with sizes $s_1, \ldots, s_m$ have to be packed. Then, $\mathcal{P}^+$ is formed by $\{(t_i, s_i)\}$. Moreover, the algorithms under consideration will regularly create placeholder items (called *virtual items*) at time points $\{\tilde{t}_1, \ldots, \tilde{t}_u\}$ with sizes $\tilde{s}_i$. These pairs $\{(\tilde{t}_i, \tilde{s}_i)\}$ form the set $\mathcal{P}^-$. Then, the goal is to replace as many placeholders by real items. So, we interpret the x-coordinates as time component, and the y-coordinates as the sizes of the items or the placeholders.

A maximum matching for $\mathcal{P}$ can be computed as follows: We process the points in $\mathcal{P}^+$ in increasing order according to their x-coordinates. Then, we try to match the point (x^+, y^+) currently processed to an unmatched point (x^-, y^-) from $\mathcal{P}^-$ with $x^- \leq x^+$ and $y^- \geq y^+$ with y^- as small as possible. If no such point (x^-, y^-) exists, then (x^+, y^+) will be left unmatched.

Moreover, it follows from this algorithm that it is also possible to obtain a maximum matching if we assume that the points in $\mathcal{P}^+ \cup \mathcal{P}^-$ arrive online in order according to their x-coordinates (we assume here for simplicity that the x-coordinates of all points in $\mathcal{P}^+ \cup \mathcal{P}^-$ are different).

2.3.1 Problems and Estimates

We are interested in variants of the problem where the point sets $\mathcal{P}^+$ and $\mathcal{P}^-$ are generated randomly in different ways. When we deal with an optimal algorithm in the random-order model for cardinality-constrained bin packing (and afterwards in the analysis of an $(1+\epsilon)$-competitive algorithm in the random-order model for the classical bin packing problem) two matching variants arise. In the first variant the plus- and minus-points are arranged alternately according to their x-coordinate, and the y-coordinates are drawn from a *common* set using sampling without replacement.

MATCHING VARIANT 1

Let $n, k \in \mathbb{N}$ with $2k \leq n$, and $\mathcal{H} = \{h_1, \ldots, h_n\}$ with $h_i = i$. Moreover, let π be a random permutation of the elements in $[n]$. We set $H_i = h_{\pi(i)}$ and then we set

$$\mathcal{P}^+ = \{(2i, H_{2i})\}_{1\leq i\leq k} \qquad \text{and}$$
$$\mathcal{P}^- = \{(2i-1, H_{2i-1})\}_{1\leq i\leq k}.$$

We will introduce more matching variants in the future. Here, the i-th matching variant will also be abbreviated by (Mi).

In the second variant the arrangement of the points with respect to the x-coordinate is according to a random permutation. Moreover, the y-coordinates of the points in $\mathcal{P}^+$ and $\mathcal{P}^-$ are drawn from two *different* sets.

MATCHING VARIANT 2

Let $n, k \in \mathbb{N}$ with $k \leq n$, $\mathcal{H}^+ = \{h_1^+, \ldots, h_n^+\}$ with $h_i^+ = 2i-1$, and $\mathcal{H}^- = \{h_1^-, \ldots, h_n^-\}$ with $h_i^- = 2i$. Moreover, let $\mathcal{L} = \{\ell_1, \ldots, \ell_{2n}\}$ with $\ell_i = 1$ for $1 \leq i \leq n$ and $\ell_i = -1$ for $n+1 \leq i \leq 2n$. For a random permutation σ of $[2n]$, let $L_i = \ell_{\sigma(i)}$. Then, we set $\mathcal{J}_\sigma^+ = \{i \in [k] : L_i = 1\}$, and $\mathcal{J}_\sigma^- = \{i \in [k] : L_i = -1\}$. Furthermore, let π^+, π^- be two independent random permutations of the elements in $[n]$, $H_i^+ = h^+_{\pi^+(i)}$, $H_i^- = h^-_{\pi^-(i)}$, and then we set

$$\mathcal{P}^+ = \{(i, H_i^+)\}_{i\in\mathcal{J}_\sigma^+} \qquad \text{and}$$
$$\mathcal{P}^- = \{(i, H_i^-)\}_{i\in\mathcal{J}_\sigma^-}.$$

The crucial point is that the number of unmatched points in a random instance $\mathcal{P}$ generated with respect to (M1) (or (M2)), is sublinear with high probability:

▶ **Lemma 25.** *Let $\mathcal{P}$ be a random instance of the matching-variant (M1) (or (M2) respectively). Then, there exist universal (that is independent of n and k) constants α, C, K such that it holds*

$$\mathbb{P}\left[U(\mathcal{P}) \geq K\sqrt{k}\log(k)^{3/4}\right] \leq C\exp\left(-\alpha\log(k)^{3/2}\right).$$

We were not able to find a reference for these two matching variants and the bound stated in the previous lemma. Therefore, we will give a formal deduction of this bound from existing ones in Section 2.3.2.

The following two variants, which are relevant for the analysis of class-constrained bin packing, can be found in [72]. At first the points in $\mathcal{P}^+$ and $\mathcal{P}^-$ are again arranged alternately with respect to the x-coordinates. The y-coordinates are distributed according to two different distributions μ and ν.

> MATCHING VARIANT 3
>
> Let $n \in \mathbb{N}$, and μ, ν be two probability measures on $[0,1]$. Let $X_1, \dots, X_n$ be independent and distributed according to μ, and $Y_1, \dots, Y_n$ be independent and distributed according to ν. Then, we set
>
> $$\mathcal{P}^+ = \{(2i, X_i)\}_{1\le i\le n} \quad \text{and}$$
> $$\mathcal{P}^- = \{(2i-1, Y_i)\}_{1\le i\le n}.$$

In the fourth variant a point is with probability 1/2 a plus-point, and a minus-point otherwise. The y-coordinates of plus- and minus-points are generated according to a common distribution μ.

> MATCHING VARIANT 4
>
> Let $n \in \mathbb{N}$, and μ be a probability measure on $[0,1]$. Let $X_1, \dots, X_n$ be independent and distributed according to μ, and $L_1, \dots, L_n$ independent random variables with $\mathbb{P}[L_i = 1] = \mathbb{P}[L_i = -1] = 1/2$. Then, we set
>
> $$\mathcal{P}^+ = \{(i, X_i) \,:\, L_i = 1,\, i \in [n]\}$$
> $$\mathcal{P}^- = \{(i, X_i) \,:\, L_i = -1,\, i \in [n]\}.$$

Here, we have the following two bounds:

▶ **Lemma 26.** *Let $\mathcal{P}$ be an instance for (M3). We define*

$$d(\mu,\nu) := \sup_{0\le t\le 1} \{\mu([t,1]) - \nu([t,1])\}.$$

Then, there exist universal constants (that is independent of n, μ and ν) α, C, K such that

$$\mathbb{P}\left[U(\mathcal{P}) \ge d(\mu,\nu)\cdot n + K\sqrt{n}\log(n)^{3/4}\right] \le C\exp\left(-\alpha\log(n)^{3/2}\right).$$

▶ **Lemma 27.** *Let $\mathcal{P}$ be an instance for (M4). Then, there exist universal constants (that is independent of μ and n) such that*

$$\mathbb{P}\left[U(\mathcal{P}) \ge K\sqrt{n}\log(n)^{3/4}\right] \le C\exp\left(-\alpha\log(n)^{3/2}\right).$$

2.3.2 Deferred Proofs

In this part we will give formal proofs of the bounds for Matching Variant 1 and 2 stated in Lemma 25. The derivation is very technical and the applied techniques are not important for the proofs given Section 3-5.

In order to prove Lemma 25 we will deal with four more matching variants: Perhaps the fundamental version of our matching variants is the following, which was studied in [59] and [71]:

> MATCHING VARIANT 5
>
> Let $X_1, \dots, X_{2n}, Y_1, \dots, Y_{2n}$ be independent, $\mathcal{U}[0,1]$-distributed random variables. Then, we set
>
> $$\mathcal{P}^+ = \{(X_i, Y_i)\}_{1\le i\le n} \quad \text{and}$$
> $$\mathcal{P}^- = \{(X_i, Y_i)\}_{n+1\le i\le 2n}.$$

In this model, with probability 1 all points have different x- and y-coordinates. Since the possible matchings only depend on the relative position of the points, this variant is equivalent to the following one:

Matching Variant 6

Let $\mathcal{H} = \{h_1, \dots, h_{2n}\}$ with $h_i = i$, and $\mathcal{L} = \{\ell_1, \dots, \ell_{2n}\}$ with $\ell_i = 1$ for $1 \leq i \leq n$ and $\ell_i = -1$ for $n+1 \leq i \leq 2n$. Moreover, let π, σ be two independent random permutations of the elements in $[2n]$. Then, we set $L_i = \ell_{\sigma(i)}$ and $H_i = h_{\pi(i)}$. Eventually, we set

$$\mathcal{P}^+ = \{(i, H_i) \,:\, L_i = 1,\, i \in [2n]\}$$
$$\mathcal{P}^- = \{(i, H_i) \,:\, L_i = -1,\, i \in [2n]\}.$$

In [59] and [71] it was shown that the following bound holds:

▶ **Lemma 28.** *Let $\mathcal{P}$ be a point set generated according to (M5) or (M6). Then, there exist universal constants α, C, K such that we have*

$$\mathbb{P}\left[U(\mathcal{P}) \geq K\sqrt{n}\log(n)^{3/4}\right] \leq C\exp\left(-\alpha\log(n)^{3/2}\right).$$

At first we will derive the bound for (M1). Starting from the bound given in Lemma 28 for (M6) we will show that essentially the same bound holds for an intermediate matching problem.

2.3.2.1 An Intermediate Matching Variant

Matching Variant 7

Let $n \in \mathbb{N}$, and $\mathcal{H} = \{h_1, \dots, h_{2n}\}$ with $h_i = i$. Moreover, let π be a random permutation of the elements in $[2n]$, and we set $H_i = h_{\pi(i)}$. Then, we set

$$\mathcal{P}^+ = \{(2i, H_{2i})\}_{1\leq i\leq n} \qquad \text{and}$$
$$\mathcal{P}^- = \{(2i-1, H_{2i-1})\}_{1\leq i\leq n}.$$

We see that an instance $\tilde{\mathcal{P}}$ for (M7) depends only on the permutations of heights, so let $\{\tilde{\mathcal{P}}_\pi\}_\pi$ denote the set of all possible realizations of the instance. (Choosing a random instance for (M7) corresponds to choosing a random element from $\{\tilde{\mathcal{P}}_\pi\}_\pi$ with respect to the uniform distribution.)

On the other hand, an instance $\mathcal{P}$ for (M6) depends on a permutation of the heights π, and a permutation of the labels σ. So the set of possible realizations is given by $\{\mathcal{P}_{\sigma,\pi}\}_{\sigma,\pi}$ (and again generating a random instance corresponds to choosing a random element from this set with respect to the uniform distribution).

In the following, we will fix a permutation σ of the labels and will construct a bijection m_σ between the sets $\{\mathcal{P}_{\sigma,\pi}\}_\pi$ and $\{\tilde{\mathcal{P}}_\pi\}_\pi$ in such a way that the number of unmatched plus-points in $m_\sigma(\mathcal{P}_{\sigma,\pi})$ is upper bounded by the number of unmatched plus-points in $\mathcal{P}_{\sigma,\pi}$ plus an additive term that depends only on σ.

For a fixed permutation σ let

$$\mathcal{J}_\sigma^+ = \{j_1^+, \dots, j_n^+\}$$

denote the indices such that $L_{j_i^+} = 1$ in $\mathcal{P}_{\sigma,\pi}$. We assume without loss of generality that $j_1^+ < \ldots < j_n^+$. Moreover, we define $\mathcal{J}_\sigma^-$ in an analogous way.

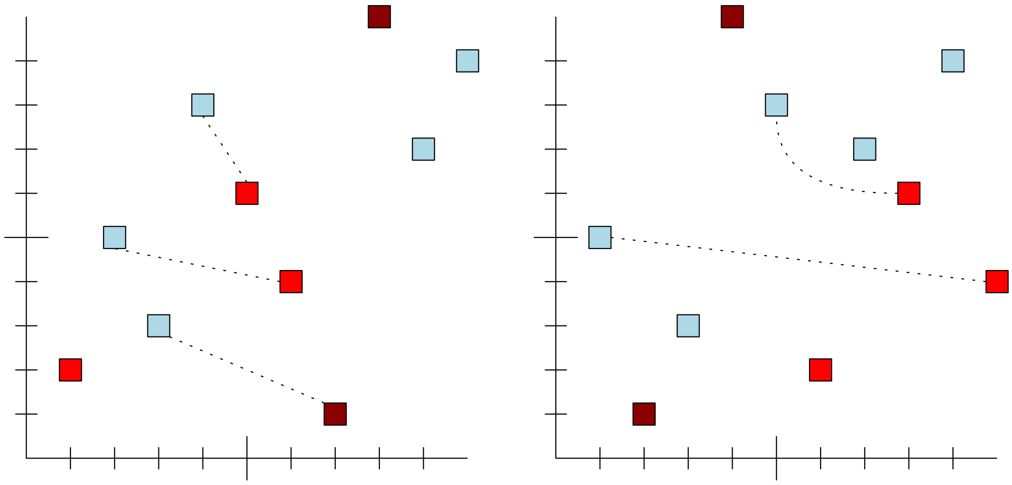

Figure 6 The mapping m_σ in the reduction from (M7) to (M6) for a fixed permutation of the labels $(1,-1,-1,-1,1,1,1,1,-1,-1)$ and a permutation π of the heights. We have $\mathcal{B}_\sigma = \{7,8\}$ and $\mathcal{G}_\sigma = \{1,5,6\}$. On the left hand side we see $\mathcal{P}_{\sigma,\pi}$ and on the right hand side $m_\sigma(\mathcal{P}_{\sigma,\pi})$. The dark-red squares indicate bad plus-points, the remaining red squares good plus-points, and the blue squares denote minus-points.

For $i \in [2n]$ we set $N_i^+ := |\{j \le i \,:\, L_j = 1\}|$ and $N_i^- := |\{j \le i \,:\, L_j = -1\}|$. Let

$$D_\sigma := \max_{1\le k\le 2n} \left(N_k^- - N_k^+\right).$$

Now, we split up the plus-points $\mathcal{P}_{\sigma,\pi}^+$ into *good* and *bad* plus-points: The good plus-points are the ones with x-coordinates in $\mathcal{G}_\sigma := \{j_1^+, \dots, j_{n-D_\sigma}^+\}$ and the bad ones with x-coordinates in $\mathcal{B}_\sigma := \{j_{n-D_\sigma+1}^+, \dots, j_n^+\}$.

We construct a bijection m_σ between the instances $\{\mathcal{P}_{\sigma,\pi}\}_\pi$ and $\{\tilde{\mathcal{P}}_\pi\}_\pi$: In order to do this, we map the minus-points in $\mathcal{P}_{\sigma,\pi}$ to the odd positions according to their order, and map the plus-points such that the bad ones begin and then the good ones follow. More formally, starting from $\mathcal{P}_{\sigma,\pi}^- = \{(j_i^-, H_{j_i^-})\}_{1\le i\le n}$ and $\mathcal{P}_{\sigma,\pi}^+ = \{(j_i^+, H_{j_i^+})\}_{1\le i\le n}$, we set

$$m_\sigma(\mathcal{P}_{\sigma,\pi}^-) = \{(2i-1, H_{j_i^-})\}_{1\le i\le n}$$
$$m_\sigma(\mathcal{P}_{\sigma,\pi}^+) = \{(2i, H_{b_i})\}_{1\le i\le |\mathcal{B}_\sigma|} \cup \{(2(|\mathcal{B}_\sigma|+i), H_{g_i})\}_{1\le i\le |\mathcal{G}_\sigma|}.$$

In Figure 6 we give a concrete example.

Let $\mathcal{P}_{\sigma,\pi,r}$ denote the restricted instance where we take only good plus-points into account, i.e., we have $\mathcal{P}_{\sigma,\pi,r}^- = \mathcal{P}_{\sigma,\pi}^-$ and $\mathcal{P}_{\sigma,\pi,r}^+ = \{(j_i^+, H_{j_i^+}, 1)\}_{1\le i\le n-D_\sigma}$.

We have constructed m_σ in such a way that every matching for $\mathcal{P}_{\pi,\sigma,r}$ is also a matching for $m_\sigma(\mathcal{P}_{\pi,\sigma})$: Since we only change the order of the points with respect to their x-coordinate, we only have to show that if we map a point from $\mathcal{P}_{\pi,\sigma,r}^+$ with x-coordinate j_a^+ to a point from $\mathcal{P}_{\pi,\sigma,r}^-$ with x-coordinate j_b^-, then the transformed x-coordinates have to maintain their relative order. Because of the monotonicity of our mapping, it suffices to show this for two consecutive points, that is, we have $j_a^+ = j_b^- + 1$. The x-coordinate j_a^+ will be mapped to $2D_\sigma + 2a$ and j_b^- to $2b-1$. We claim that $D_\sigma \ge b-a$, which would show the statement. Since we look at two consecutive points it is $j_a^+ = a + N_{j_b^-}^-$ and $j_b^- = b + N_{j_b^-}^+$.

It follows that

$$b - a = \left(j_b^- - N_{j_b^-}^+\right) - \left(j_a^+ - N_{j_b^-}^-\right) = j_b^- - j_a^+ + N_{j_b^-}^- - N_{j_b^-}^+ < N_{j_b^-}^- - N_{j_b^-}^+ \le D_\sigma.$$

Since we can adapt every matching for $\mathcal{P}_{\sigma,\pi,r}$ to a matching for $m_\sigma(\mathcal{P}_{\sigma,\pi})$ with same cardinality we have $U(m_\sigma(\mathcal{P}_{\sigma,\pi})) \leq U(\mathcal{P}_{\sigma,\pi,r})$. Moreover, we have $U(\mathcal{P}_{\pi,\sigma,r}) \leq U(\mathcal{P}_{\pi,\sigma}) + D_\sigma$, and hence we obtain

$$U(m_\sigma(\mathcal{P}_{\sigma,\pi})) \leq U(\mathcal{P}_{\sigma,\pi}) + D_\sigma.$$

Using this inequality, we see that

$$\begin{aligned}
&\mathbb{P}\left[U(\tilde{\mathcal{P}}_\pi) \geq 2K\sqrt{n}\log(n)^{3/4}\right] \\
=&\mathbb{P}\left[U(m_\sigma(\mathcal{P}_{\sigma,\pi})) \geq 2K\sqrt{n}\log(n)^{3/4}\right] \\
\leq&\mathbb{P}\left[U(\mathcal{P}_{\sigma,\pi}) + D_\sigma \geq 2K\sqrt{n}\log(n)^{3/4}\right] \\
\leq&\mathbb{P}\left[U(\mathcal{P}_{\sigma,\pi}) \geq K\sqrt{n}\log(n)^{3/4}\right] + \mathbb{P}\left[D_\sigma \geq K\sqrt{n}\log(n)^{3/4}\right].
\end{aligned}$$

Using Lemma 28 and the following lemma, then yields the bound for (M7).

▶ **Lemma 29.** *Let K be an arbitrary constant, then there exist constants α', C' such that we have*

$$\mathbb{P}\left[D_\sigma \geq K\sqrt{n}\log(n)^{3/4}\right] \leq C' \exp\left(-\alpha' \log(n)^{3/2}\right).$$

Proof of Lemma 29. Let $S_k := \sum_{i=1}^{k} L_i$. We observe that D_σ is the same as $-\min_{1\leq k\leq 2n} S_k$. Applying Proposition 14 we see that

$$\mathbb{P}\left[\max_{1\leq k\leq 2n} \|S_k\| > K\sqrt{n}\log(n)^{3/4}\right] \leq c \cdot \mathbb{P}\left[\|S_n\| > \frac{K}{c} \cdot \sqrt{n}\log(n)^{3/4}\right],$$

with an universal constant c. Hence, applying Hoeffding's inequality, we obtain

$$\mathbb{P}\left[\max_{1\leq k\leq 2n} \|S_k\| > K\sqrt{n}\log(n)^{3/4}\right] \leq 2c \cdot \exp\left(-\frac{K^2}{2c^2} \cdot \log(n)^{3/2}\right).$$

◀

2.3.2.2 Deriving the Bound for (M1) from (M7)

Now we want to relate (M1) and (M7). We observe that we can generate a random instance for (M1) as follows: In the first step we draw a random subset H of $\mathcal{H}$ of size $2k$ using sampling without replacement. Then, we assign the heights from H to the coordinates using a random permutation of the elements in $[2k]$. We observe that – since only the relative positions of the heights matter – we can assume without loss of generality that the random subset H is given by $[2k]$. It follows that for each fixed H we have an equivalent instance to (M7) with $n = k$. Thus, we obtain exactly the same bound as for (M7) and this shows Lemma 25 for (M1).

2.3.2.3 Another Intermediate Matching Variant

In order to show the bound for (M2) we introduce another intermediate matching variant. Again, starting from the bound given Lemma 28 for (M6) we will show that essentially the same bound holds for the following variation:

MATCHING VARIANT 8

Let $n \in \mathbb{N}$, $\mathcal{H}^+ = \{h_1^+, \ldots, h_n^+\}$ with $h_i^+ = 2i-1$, and $\mathcal{H}^- = \{h_1^-, \ldots, h_n^-\}$ with $h_i^- = 2i$. Let π^+, π^- be two independent random permutations of the elements in $[n]$. Then, we set $H_i^+ = h^+_{\pi^+(i)}$ and $H_i^- = h^-_{\pi^-(i)}$. Finally, we set

$$\mathcal{P}^+ = \{(2i, H_i^+)\}_{1\le i\le n} \qquad \text{and}$$
$$\mathcal{P}^- = \{(2i-1, H_i^-)\}_{1\le i\le n}.$$

We want to relate this matching variant to (M7). We can generate a random instance $\mathcal{P}$ for (M7) as follows: At first we choose a random subset (by sampling without replacement n elements from $[2n]$) $\{h_1^+, \ldots, h_n^+\} =: H^+ \subseteq [2n]$ of the heights and set $H^- := [2n]\setminus H^+$. Afterwards we generate two independent random permutations π^+, π^- of the elements in $[n]$ and set

$$\mathcal{P}^+ := \{(2i, h^+_{\pi^+(i)})\}_{1\le i\le n}$$
$$\mathcal{P}^- := \{(2i-1, h^-_{\pi^-(i)})\}_{1\le i\le n}.$$

We notice that there are $\binom{2n}{n} = (2n)!/(n!)^2$ many possible random subsets H^+ of the heights of the plus-points. Now, for fixed H^+ we want to relate the $(n!)^2$ instances $\mathcal{P}_{H^+,\pi^+,\pi^-}$ for (M7) with the $(n!)^2$ instances $\tilde{\mathcal{P}}_{\pi^+,\pi^-}$ for (M8). That is, we construct a mapping which preserves the maximum matching apart from an error that is of order $\sqrt{n}\log(n)^{3/4}$ with high probability.

Let $f^+ : H^+ \to \{1, 3, \ldots, 2n-1\}$ denote a bijection that maximizes the number of $i \in H^+$ such that $f^+(i) \le i$. Moreover, let $f^- : H^- \to \{2, 4, \ldots, 2n\}$ denote a bijection that maximizes the number of $i \in H^-$ such that $f^-(i) \ge i$. Now, we want to give a bijection based on f^+ and f^- between the instances $\mathcal{P}_{H^+,\pi^+,\pi^-}$ for (M7) with a fixed set H^+ and the instances for (M8). In order to do this, we map the set $\mathcal{P}^+ = \{(2i, h^+_{\pi^+(i)})\}_{i\in[n]}$ to $\tilde{\mathcal{P}}^+ = \{(2i, f^+(h^+_{\pi^+(i)}))\}_{i\in[n]}$ and $\mathcal{P}^- = \{(2i-1, h^+_{\pi^-(i)})\}_{i\in[n]}$ to $\tilde{\mathcal{P}}^- = \{(2i-1, f^-(h^-_{\pi^-(i)}))\}_{i\in[n]}$. Let us call this mapping m_{H^+}. m_{H^+} is constructed in such a way that we can adopt the matching for $\mathcal{P}_{H^+,\pi^+,\pi^-}$ to $m_{H^+}(\mathcal{P}_{H^+,\pi^+,\pi^-})$ apart from the points with heights such that $\{i \in H^+ : f^+(i) > i\}$ or $\{i \in H^- : f^-(i) < i\}$. An example for this mapping is given in Figure 7.

Let $D_{H^+} := |\{i \in H^+ : f^+(i) > i\}| + |\{i \in H^- : f^-(i) < i\}|$. Then, we have

$$U(m_{H^+}(\mathcal{P}_{H^+,\pi^+,\pi^-})) \le U(\mathcal{P}_{H^+,\pi^+,\pi^-}) + D_{H^+}.$$

It follows that

$$\begin{aligned}
&\mathbb{P}\left[U(\tilde{\mathcal{P}}_{\pi^+,\pi^-}) \ge 2K\sqrt{n}\log(n)^{3/4}\right]\\
=&\mathbb{P}\left[U(m_{H^+}(\mathcal{P}_{H^+,\pi^+,\pi^-})) \ge 2K\sqrt{n}\log(n)^{3/4}\right]\\
\le&\mathbb{P}\left[U(\mathcal{P}_{H^+,\pi^+,\pi^-}) + D_{H^+} \ge 2K\sqrt{n}\log(n)^{3/4}\right]\\
\le&\mathbb{P}\left[U(\mathcal{P}_{H^+,\pi^+,\pi^-}) \ge K\sqrt{n}\log(n)^{3/4}\right] + \mathbb{P}\left[D_{H^+} \ge K\sqrt{n}\log(n)^{3/4}\right].
\end{aligned}$$

Combining the bound for (M7) and the next lemma we obtain the bound for (M8).

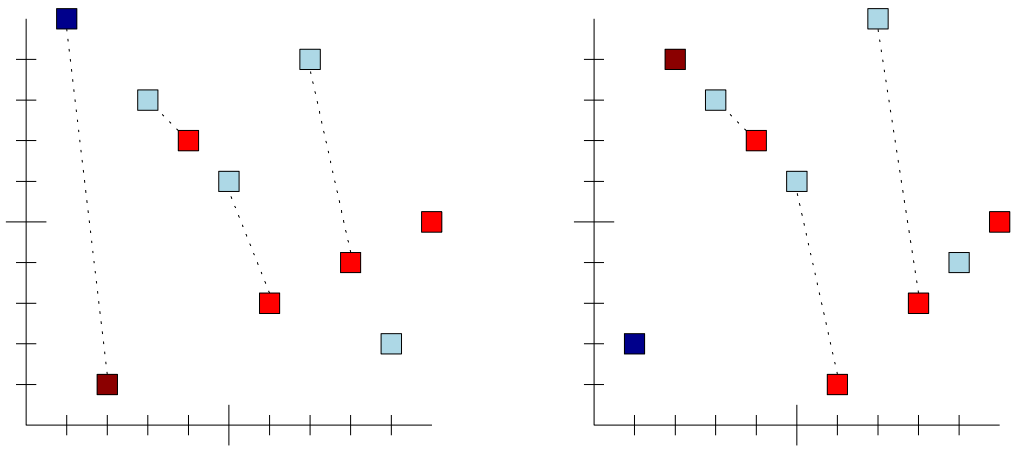

Figure 7 The reduction from an instance for (M8) to (M7) with $n = 5$. The random subset H^+ is given by $\{1, 3, 4, 5, 7\}$. We set $f^+(3) = 1$, $f^+(4) = 3$, $f^+(5) = 5$, $f^+(7) = 7$ and $f^+(1) = 9$. Moreover, we have $f^-(2) = 4$, $f^-(6) = 6$, $f^-(8) = 8$, $f^-(9) = 10$, and $f^-(10) = 2$. Plus-points are given by red squares. Dark-red squares denote plus-points with a height i such that $f^+(i) > i$. Dark-blue squares are defined analogously.

▶ **Lemma 30.** *Let K be an arbitrary constant. Then, there exist constants α', C' such that*

$$\mathbb{P}\left[D_{H^+} \geq K\sqrt{n}\log(n)^{3/4}\right] \leq C' \exp\left(-\alpha' \log(n)^{3/2}\right).$$

Proof. For a random drawn set H^+ and $H^- := [2n]\backslash H^+$, and $1 \leq i \leq 2n$ let

$$X_i = \begin{cases} 1, & i \in H^+ \\ -1, & i \in H^-. \end{cases}$$

Then, we notice that we can bound $|\{i \in H^+ : f^+(i) > i\}|$ and $|\{i \in H^- : f^-(i) < i\}|$ in terms of $\max_{1 \leq j \leq 2n} \left|\sum_{i=1}^{j} X_i\right|$. So it remains to give bounds for the maximum of the partial sums. This can be done applying at first the maximum inequality for sampling without replacement (Proposition 14) and afterwards Hoeffding's inequality. ◀

2.3.2.4 Deriving the Bound for (M2) from (M8)

This deduction is more tedious as we have to map points vertically and horizontally.

We can generate a random instance $\tilde{\mathcal{P}}$ for (M2) in several steps as follows: At first we determine the order of the labels. Let $\mathcal{L} = \{\ell_1, \dots, \ell_{2n}\}$ with $\ell_i = 1$ for $1 \leq i \leq n$ and $\ell_i = -1$ for $n+1 \leq i \leq 2n$. Let σ be a random permutation of the elements in $[2n]$. Then, we set $L_i = \ell_{\sigma(i)}$. Let $N_\sigma^+ = |\{i \leq k : L_i = 1\}|$, $N_\sigma^- = |\{i \leq k : L_i = -1\}|$ and $M_\sigma = \min\{N_\sigma^+, N_\sigma^-\}$.

Let $\mathcal{J}_\sigma^+ = \{j_1^+, \dots, j_{N_\sigma^+}^+\}$ denote the x-coordinates of the plus-points and let $\mathcal{J}_\sigma^- = \{j_1^-, \dots, j_{N_\sigma^-}^-\}$ denote the x-coordinates of the minus-points. (We assume again that $j_1^+ < \dots < j_{N_\sigma^+}^+$ and $j_1^- < \dots < j_{N_\sigma^-}^-$.) Furthermore, we set $\check{\mathcal{J}}_\sigma^+ = \{j_1^+, \dots, j_{M_\sigma}^+\}$ and $\check{\mathcal{J}}_\sigma^- = \{j_1^-, \dots, j_{M_\sigma}^-\}$.

Usually, we will have either $\mathcal{J}_\sigma^+ \backslash \check{\mathcal{J}}_\sigma^+ \neq \emptyset$ or $\mathcal{J}_\sigma^- \backslash \check{\mathcal{J}}_\sigma^- \neq \emptyset$. We call the points with x-coordinates in the non-empty set *surplus-points*.

In the next step we generate the set of heights H^s for the surplus-points by drawing $k - 2M_\sigma$ many elements using sampling without replacement from $[2n]$. Furthermore, a random permutation π^s of the elements in $[k - 2M_\sigma]$ is used to assign the heights in H^s to the surplus-points.

Then, we determine the set of heights for plus-points H^+ by drawing M_σ many elements with sampling without replacement from $[2n]\backslash H^s$ and afterwards we generate H^- by drawing M_σ many elements with sampling without replacement from $[2n]\backslash(H^s \cup H^+)$.

Finally, we generate two independent random permutations π^+, π^- of the elements in $[M_\sigma]$ to assign the heights of plus-points (minus-points) to the positions.

We see that $\tilde{\mathcal{P}}$ depends on $\sigma, H^s, \pi^s, H^+, H^-, \pi^+$ and π^-, so we will write in the following $\tilde{\mathcal{P}} = \tilde{\mathcal{P}}_{\sigma,H^s,\pi^s,H^+,H^-,\pi^+,\pi^-}$. An instance $\mathcal{P}$ for (M8) with $n = M_\sigma$ just depends on two random permutations π^+, π^-, which assign the heights to the positions, so we write $\mathcal{P}_{\pi^+,\pi^-}$.

Assume that $\sigma, H^s, \pi^s, H^+, H^-$ are fixed. In the following we want to construct a bijection m_{σ,H^+,H^-} between $\{\mathcal{P}_{\pi^+,\pi^-}\}_{\pi^+,\pi^-}$ and $\{\tilde{\mathcal{P}}_{\sigma,H^s,\pi^s,H^+,H^-,\pi^+,\pi^-}\}_{\pi^+,\pi^-}$. The aim is to choose m_{σ,H^+,H^-} in such a way that we can adapt the matching for $\mathcal{P}_{\pi^+,\pi^-}$ to $m_{\sigma,H^+,H^-}(\mathcal{P}_{\pi^+,\pi^-})$ for most of the points.

Now we describe the construction of the bijection in several steps. An example is given in Figure 8.

1. At first we transform x-coordinates in a similar way as done in the reduction from (M6) to (M7): Let $D_\sigma := \max_{1\le i\le k} L_i$. Outgoing from $\mathcal{P}^+_{\pi^+,\pi^-} = \{(2i, H_i^+)\}_{1\le i\le M_\sigma}$ and $\mathcal{P}^-_{\pi^+,\pi^-} = \{(2i-1, H_i^-)\}_{1\le i\le M_\sigma}$ we map them to

$$\left\{(j^+_{1+((D_\sigma+i-1) \bmod M_\sigma)}, H_i^+)\right\}_{1\le i\le M_\sigma} \quad \text{and} \quad \left\{(j_i^-, H_i^-)\right\}_{1\le i\le M_\sigma}. \tag{3}$$

We notice that apart from D_σ many plus-points, the matching for $\mathcal{P}_{\pi^+,\pi^-}$ can be preserved.

2. The next step is the transformation of heights: We want to map as many plus-points as possible to a greater height, and as many minus-points to a lower height. Given a set $H^+ \subseteq [2M_\sigma]$ with $|H^+| = M_\sigma$ let $f^+ : \{2,4,\dots,2M_\sigma\} \to H^+$ denote a bijection that maximizes $|\{e \in \{2,4,\dots,2M_\sigma\} : f^+(e) \ge e\}|$ and respectively for $H^- \subseteq [2M_\sigma]$ with $|H^-| = M_\sigma$ let $f^- : \{1,3,\dots,2M_\sigma - 1\} \to H^-$ denote a bijection that maximizes $|\{e \in \{1,3,\dots,2M_\sigma-1\} : f^-(e) \le e\}|$. Moreover, let

$$E_{H^+} := \left|\left\{e \in \{2,4,\dots,2M_\sigma\} : f^+(e) < e\right\}\right|$$
$$E_{H^-} := \left|\left\{e \in \{1,3,\dots,2M_\sigma-1\} : f^-(e) > e\right\}\right|.$$

Then, we map the set given in (3) to

$$\left\{(j^+_{1+((D_\sigma+i-1) \bmod M_\sigma)}, f^+(H_i^+))\right\}_{1\le i\le M_\sigma} \quad \text{and} \quad \left\{(j_i^-, f^-(H_i^-))\right\}_{1\le i\le M_\sigma}. \tag{4}$$

We see that we lose at most $E_{H^+} + E_{H^-}$ many matched points by this transformation.

3. Until now we have assumed that $H^+ = \{2,4,\dots,2M_\sigma\}$ and $H^- = \{1,3,\dots,2M_\sigma-1\}$. If this is not the case, then we can find a mapping g from $[2M_\sigma]$ to $H^+ \cup H^-$ that preserves the relative order between plus- and minus-points such that still each matching can be adapted.

We see that we have $U(m_{\sigma,H^+,H^-}(\mathcal{P}_{\pi^+,\pi^-})) \le U(\mathcal{P}_{\pi^+,\pi^-}) + (k - 2M_\sigma) + D_\sigma + E_{H^+} + E_{H^-}$. Hence, we obtain

$$\begin{aligned}&\mathbb{P}\left[\tilde{\mathcal{P}}_{\sigma,H^s,\pi^s,H^+,H^-,\pi^+,\pi^-} \ge 5K\sqrt{k}\log(k)^{3/4}\right]\\ =&\mathbb{P}\left[U(m_{\sigma,H^+,H^-}(\mathcal{P}_{\pi^+,\pi^-})) \ge 5K\sqrt{k}\log(k)^{3/4}\right]\end{aligned}$$

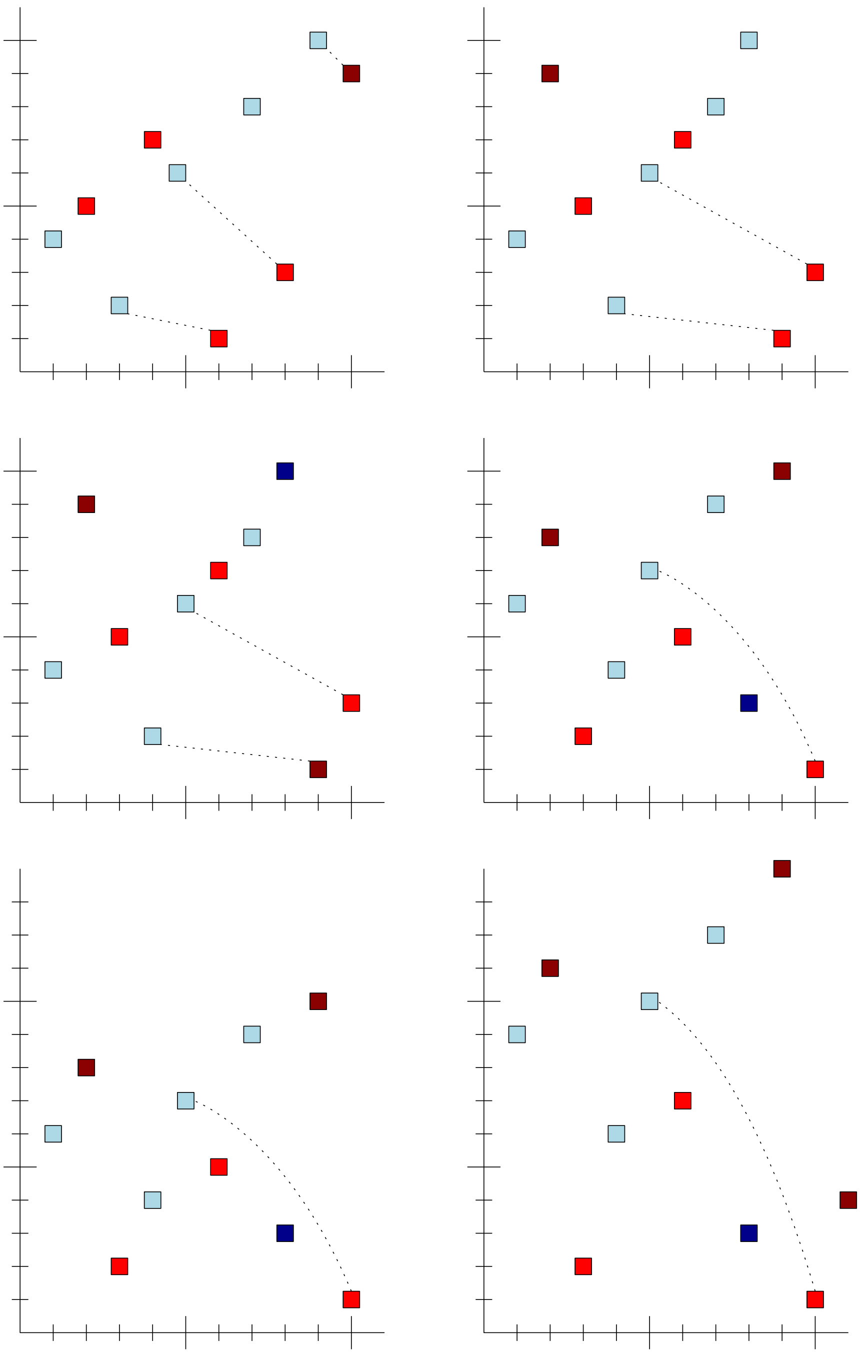

Figure 8 Assume we have an instance for $\tilde{\mathcal{P}}_{\sigma,H^s,\pi^s,H^+,H^-,\pi^+,\pi^-}$ with $(-1,1,1,-1,-1,1,-1,-1,1,1,1)$ as order of the labels, $H^s = \{4\}$, $H^+ = \{1,2,7,11,14\}$ and $H^- = \{3,6,9,10,12\}$. We illustrate the construction of m_{σ,H^+,H^-} from an instance for (M8) to $\tilde{\mathcal{P}}_{\sigma,H^s,\pi^s,H^+,H^-,\pi^+,\pi^-}$ with fixed σ, H^s, H^+, H^-. In the first row we show the transformation of the x-coordinates. Plus-points that possibly cannot be matched after this transformation are colored in dark-red. In the second row we transform the heights with $f^+(1) = 10$, $f^+(3) = 1$, $f^+(5) = 2$, $f^+(7) = 5$, and $f^+(9) = 8$. Moreover, let $f^-(2) = 4$, $f^-(4) = 6$, $f^-(6) = 7$, $f^-(8) = 9$, and $f^-(10) = 3$. New points that possibly could not be taken into account in the matching are colored in dark-red or dark-blue. In the last row we illustrate the map g with $g(1) = 1$, $g(2) = 2$, $g(3) = 3$, $g(4) = 6$, $g(5) = 7$, $g(6) = 9$, $g(7) = 10$, $g(8) = 11$, $g(9) = 12$, and $g(10) = 14$.

$$
\begin{aligned}
\leq & \mathbb{P}\left[U(\mathcal{P}_{\pi^+,\pi^-}) + (k - 2M_\sigma) + D_\sigma + E_{H^+} + E_{H^-} \geq 5K\sqrt{k}\log(k)^{3/4}\right] \\
\leq & \mathbb{P}\left[U(\mathcal{P}_{\pi^+,\pi^-}) \geq K\sqrt{k}\log(k)^{3/4}\right] + \mathbb{P}\left[(k - 2M_\sigma) \geq K\sqrt{k}\log(k)^{3/4}\right] \\
& \quad + \mathbb{P}\left[D_\sigma \geq K\sqrt{k}\log(k)^{3/4}\right] + \mathbb{P}\left[E_{H^+} \geq K\sqrt{k}\log(k)^{3/4}\right] \\
& \qquad + \mathbb{P}\left[E_{H^-} \geq K\sqrt{k}\log(k)^{3/4}\right].
\end{aligned}
$$

Applying the same estimates and concentration bounds as in the previous reductions then yields the result.

2.4 Useful Facts about Probabilistic Performance Measures

The following lemma shows – under mild conditions – the existence of an alternative representation of the average performance ratio:

▶ **Lemma 31.** *Let A be an online algorithm for one of the problems we deal with in this thesis. If $\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] \in \Theta(n)$, and the competitive ratio of A is bounded from above by a constant C, then it follows that*

$$
\mathrm{APR}(A, F) = \limsup_{n\to\infty} \frac{\mathbb{E}\left[A(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}.
$$

▶ Remark. We notice that the competitive ratio of online algorithms for classical, cardinality-constrained and class-constrained bin covering is automatically bounded from above by 1.

Proof of Lemma 31. Using the Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
& \left|\mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] - \frac{\mathbb{E}\left[A(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}\right| \\
= & \left|\mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] - \mathbb{E}\left[\frac{A(I_n^F)}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}\right]\right| \\
= & \left|\mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)} \cdot \left(1 - \frac{\mathrm{OPT}(I_n^F)}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}\right)\right]\right| \\
= & \left|\mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)} \cdot \frac{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] - \mathrm{OPT}(I_n^F)}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}\right]\right| \\
\leq & \mathbb{E}\left[\left|\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right| \cdot \left|\frac{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] - \mathrm{OPT}(I_n^F)}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}\right|\right] \\
\leq & \mathbb{E}\left[\left|\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right|^2\right]^{1/2} \cdot \frac{\mathbb{E}\left[\left|\mathrm{OPT}(I_n^F) - \mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]\right|^2\right]^{1/2}}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \\
& \qquad \leq C \cdot \frac{\mathbb{E}\left[\left|\mathrm{OPT}(I_n^F) - \mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]\right|^2\right]^{1/2}}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]}.
\end{aligned}
$$

For the mentioned problems the function OPT satisfied the bounded difference property with $c_i = 1$ for $i \in \{1, \ldots, n\}$. Therefore, if we apply the bounded difference inequality (Proposition 10) we see that

$$
\mathbb{P}\left[\left|\mathrm{OPT}(I_n^F) - \mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]\right| \geq \sqrt{n\log(n)}\right] \leq \frac{2}{n^2}.
$$

Moreover, we have $|\operatorname{OPT}(I_n^F) - \mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]| \leq n$. Hence, we have

$$\mathbb{E}\left[\left|\operatorname{OPT}(I_n^F) - \mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]\right|^2\right]^{1/2} \leq (2 + n\log(n))^{1/2}\,.$$

Since we assume that $\mathbb{E}\left[\operatorname{OPT}(I_n^F)\right] \in \Theta(n)$, we see that the difference vanishes in the limit. ◀

Furthermore, we show that the adversary in the random-order model is more powerful than in the case of items that are drawn independently and identically distributed. Therefore, lower/upper bounds carry over between these two models. This connection was mentioned in [54]. A formalization was given in [42].

▶ **Lemma 32.** *Let A be an online algorithm for classical, cardinality-constrained or class-constrained bin packing. Let F be a corresponding distribution on a finite or possibly countably infinite set of items $\mathcal{I}$. Then, it holds*

$$\operatorname{APR}(A, F) \leq \operatorname{RR}(A).$$

In case of classical or class-constrained bin covering the reversed inequality is true.

Proof. Assume we are given a distribution F. Set $L_n = \{L = (a_1, \ldots, a_n) : \mathbb{P}\left[I_n^F = L\right] > 0\}$. Then there exists a set of lists $\mathcal{L}_n$, such that $L_n = \dot{\bigcup}_{H \in \mathcal{L}_n} \{L : \exists\, \sigma \text{ s.t. } L = H^\sigma\}$. Using the inequality $\left(\sum_{i=1}^n b_i\right) / \left(\sum_{i=1}^n c_i\right) \leq \max_{1 \leq i \leq n} b_i / c_i$, it follows that

$$\mathbb{E}\left[\frac{A(I_n^F)}{\operatorname{OPT}(I_n^F)}\right] \leq \max_{H \in \mathcal{L}_n} \mathbb{E}_\sigma\left[\frac{A(H^\sigma)}{\operatorname{OPT}(H^\sigma)}\right] = \max_{H \in \mathcal{L}_n} \frac{\mathbb{E}_\sigma\left[A(H^\sigma)\right]}{\operatorname{OPT}(H)}.$$

For covering problems we proceed similarly, but use the inequality $\left(\sum_{i=1}^n b_i\right) / \left(\sum_{i=1}^n c_i\right) \geq \min_{1 \leq i \leq n} b_i / c_i$. ◀

3 Complexity of Bin Packing Variants with respect to Probabilistic Performance Measures

3.1 Results

	lower bound	upper bound
CR(A)	≈ 1.542 [9]	≈ 1.5783 [8]
RR(A)	?	1.5 [54]
APR($A,\mathcal{D}$)	1	1 [72]

	lower bound	upper bound
CR(A)	≈ 1.542 [9]	≈ 1.5783 [8]
RR(A)	**1**	$\mathbf{1+\epsilon}$
APR($A,\mathcal{D}$)	1	1 [72]

Figure 9 Comparison between previously known bounds for classical bin packing on the left hand side and our results on the right hand side.

	lower bound	upper bound
CR(A)	2 [7]	2 [6]
RR(A)	?	?
APR($A,\mathcal{D}$)	?	?

	lower bound	upper bound
CR(A)	2 [7]	2 [6]
RR(A)	**1**	**1**
APR($A,\mathcal{D}$)	**1**	**1**

Figure 10 Comparison between previously known bounds for cardinality-constrained bin packing on the left hand side and our results on the right hand side.

	lower bound	upper bound
CR(A)	2 [7]	2.635 [38]
RR(A)	?	?
APR($A,\mathcal{D}$)	?	?

	lower bound	upper bound
CR(A)	2 [7]	2.635 [38]
RR(A)	$\mathbf{1.\bar{1}}$	?
APR($A,\mathcal{D}$)	**1**	**1**

Figure 11 Comparison between previously known bounds for class-constrained bin packing on the left hand side and our results on the right hand side.

In this section we study the complexity of the three introduced online bin packing variants with respect to probabilistic performance measures. That is, we are interested in the question: Is there an algorithm with average performance ratio or random-order ratio equal to one for this bin packing variant, or can we find a non-trivial lower bound?

This is motivated by a result given by Rhee and Talagrand in [72] which shows that there exists a randomized algorithm for online classical bin packing with average performance ratio equal to one for the set of all possible distributions on $(0,1]$, and the lower bounds for the competitive ratio given for example in [9]. A summary of parts of our results in comparison with existing bounds in literature is given in Figure 9, 10 and 11.

To be more precise Rhee and Talagrand [72] show that there exists a randomized algorithm A and a universal constant K (which is independent of the considered distribution F) such that with high probability the following holds:

$$A(I_n^F) \leq \mathrm{OPT}(I_n^F) + K\sqrt{n}\log(n)^{3/4}. \tag{5}$$

The idea of the proof is the following: The input is partitioned into phases, where the length of the phases increases. At the beginning of a new phase an optimal packing of all seen items is computed. As we are dealing with online algorithms (and especially with the theoretical complexity) we do not care about the runtime of the algorithm. We call this computed packing the *model-packing*. Then, the algorithm tries to pack new arriving items similar to the model-packing. In order to do this, *virtual* items are introduced, which serve as placeholders for the actually arriving items. This procedure is non-trivial and is based on a clever application of the upright matching problem.

The first result we give is about cardinality-constrained bin packing:

▶ **Proposition 33.** *For online cardinality-constrained bin packing with parameter k there exists a randomized algorithm A with* $\mathrm{RR}(A) = 1$.

The proof of this result is similar to the one given in [72]. We observe that it is possible to replace concentration inequalities and upright matching results used by appropriate versions for the case of drawing items according to sampling without replacement. So the proof is a nice starting point for getting familiar with the ideas by Rhee and Talagrand.

The next step is to consider classical bin packing: When we take a closer look at (5) we observe that the error term (that is the number of bins the algorithm needs additionally in comparison to the optimal algorithm) is sublinear in the *number* of items. This does not pose a problem in the case that items are sampled independently and identically distributed since then the size of an optimal solution grows linearly in the number of items. But if we deal with the random-order model it is possible that the number of items grows much faster than the size of the optimal solution. So in this case the error term cannot be neglected.

Unfortunately, we are not able to give an algorithm with random-order ratio equal to one as asked in [54], but we rule out the existence of a non-trivial lower bound.

▶ **Theorem 34.** *For online classical bin packing there exists for every ϵ greater than zero a randomized algorithm A_ϵ with* $\mathrm{RR}(A_\epsilon) \leq 1 + \epsilon$.

To show the result we partition the items into large and small items depending on ϵ. Then, the large items can be packed again using the approach given by Rhee and Talagrand. If possible, small items are packed into empty space created by virtual items, whose bins are not completely filled in the model-packing.

The proof of this result is very technical and relies on repeated applications of Bernstein's inequality to give bounds that depend on the size of the optimal solution and not on the number of items, and random walk arguments to control the number of bins opened for small items.

Now, we will show that such a statement is only possible if the adversary has no control over the order of an (arbitrary small) fraction of the input. We will briefly describe the *partial-permutations model*, which was introduced by Banderier, Beier and Mehlhorn in [11]. The partial-permutations model can be understood as some kind of smoothed analysis on the *order* of the items. This concept was revisited for example in [63].

The model is the following: Let $p \in [0, 1]$ be a smoothing parameter. Then, an adversary chooses an instance $I = (a_1, \ldots, a_n)$ for the bin packing problem. Afterwards, we generate a random subset $\mathcal{H}$ of $[n]$ by selecting each element in $[n]$ independently with probability p. Finally, we permute all elements with index in $\mathcal{H}$ randomly. We see that if $p = 1$ we obtain a random permutation of *all* elements in I and if $p = 0$ the adversary is able to give a worst-case instance. So, p interpolates between both cases.

It suggests itself to generalize the random-order ratio in the following way:

$$\mathrm{RR}_p(A) := \limsup_{m\to\infty} \sup_{I\,:\,\mathrm{OPT}(I)\geq m} \frac{\mathbb{E}\left[A(I^{p,\sigma})\right]}{\mathrm{OPT}(I)},$$

where $I^{p,\sigma}$ denotes a random instance obtained from I using the procedure described above. It follows that we have $\mathrm{RR}(A) = \mathrm{RR}_1(A)$.

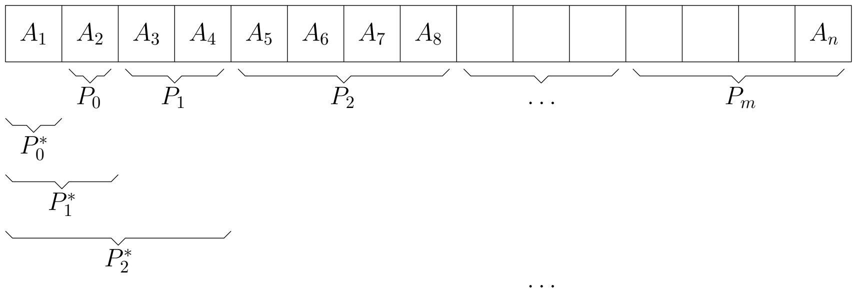

Figure 12 The partition of the input into phases.

▶ **Proposition 35.** *For online classical bin packing there exists for every $p \in [0, 1)$ an ϵ_p greater than zero such that for all deterministic online algorithms A it holds* $\mathrm{RR}_p(A) \geq 1 + \epsilon_p$.

The statement here is only given for deterministic algorithms but it is straightforward to show the same result for randomized algorithms.

Finally, we consider class-constrained bin packing. Regarding the average performance ratio we can show again that there exists a randomized algorithm which is optimal.

▶ **Proposition 36.** *Consider online class-constrained bin packing with parameter k, and let $\mathcal{D}$ denote the set of all distributions on $[0, 1] \times \mathbb{N}$. Then, there exists a randomized algorithm A with* $\mathrm{APR}(A, \mathcal{D}) = 1$.

However, the error term of our proposed algorithm can be much larger than in (5). This is especially the case if we consider distributions, for which the marginal distribution on the colors is heavy-tailed, that is we will draw items of lots of different colors. In the analysis we will apply the ideas of Rhee and Talagrand to items of an arbitrarily large but fixed subset of the colors.

This result is complemented by a non-trivial lower bound for online algorithms for class-constrained bin packing with respect to the random-order ratio:

▶ **Proposition 37.** *Consider online class-constrained bin packing with parameter k equal to 2. Then, for all deterministic online algorithms A it holds that* $\mathrm{RR}(A) \geq 10/9$.

So we see that the complexity of this problem differs if we consider the setting of items sampled with and without replacement. The statement here is again only given for deterministic algorithms, but there are only small adjustments necessary to show the same result for randomized algorithms.

3.2 Deferred Proofs

3.2.1 Existence of an Optimal Algorithm for Cardinality-constrained Bin Packing with respect to the Random-order Ratio

Here we will prove:

▶ **Proposition 33.** *For online cardinality-constrained bin packing with parameter k there exists a randomized algorithm A with* $\mathrm{RR}(A) = 1$.

3.2.1.1 Description of the Algorithm and High-Level Proof

Let $\mathcal{I} = \{a_1, \ldots, a_n\}$ be the multiset of items under consideration. Since we are dealing with cardinality-constrained bin packing it makes sense to allow that an item has size zero. The random instance we are dealing with is then given by $I^\sigma = (A_1, \dots, A_n)$ with $A_i = a_{\sigma(i)}$ where σ is a random permutation (that is random with respect to the uniform distribution on the set of all permutations). Sometimes we will call the items A_i from the input instance I^σ *real* items.

The algorithm packs item A_1 separately. Then, the sequence of items is divided into phases $P_0, \dots, P_m$. Here, the i-th phase covers the items $(A_{2^i+1}, \dots, A_{\min\{2^{i+1}, n\}})$. m is the index of the last phase, that is, m is the smallest integer such that $2^{m+1} \geq n$. The choice of the phases is illustrated in Figure 12.

For phase P_i let P_i^* denote the multiset of all items drawn before phase P_i starts, that is, it covers the items $\{A_1, \dots, A_{2^i}\}$. Moreover, $|P_i|$ and $|P_i^*|$ denote the lengths, that is the number of items in phase P_i or the number of items previously seen. The algorithm packs the items from each phase separately. In order to do this, it proceeds in the following way: At the beginning of phase P_i, that is, before item A_{2^i+1} arrives we construct an optimal packing of all items from P_i^*. We call this packing the *model-packing* $\mathcal{M}(P_i^*)$. Furthermore, let $\mathcal{Z}_i = \{A_1, \dots, A_{2^i}\}$ denote the multiset of all items we have already seen. At the time point item A_j with $j \in \{2^i + 1, \dots, \min\{2^{i+1}, n\}\}$ arrives, we create a *virtual* item V_j as follows: We draw a random item Y from $\mathcal{Z}_i$ and remove it from the set, that is, $\mathcal{Z}_i := \mathcal{Z}_i \backslash \{Y\}$. Then, we set $V_j = Y$.

To pack this virtual item we use an auxiliary algorithm A' which is based on $\mathcal{M}(P_i^*)$. This algorithm will be described in Section 3.2.1.2. Afterwards, we try to replace the smallest virtual item V with $V \geq A_j$ in our packing by A_j. If that is not possible, then we put A_j into a new bin.

Later on we will see that with high probability there will be at most $K\sqrt{|P_i|}\log(|P_i|)^{3/4}$ many real items in phase P_i that are not able to replace an appropriate item in the model-packing. As the value of the optimum solution in cardinality-constrained bin packing grows linearly in the number of items that would be sufficient for an asymptotic optimal algorithm. So what is the reason for introducing virtual items? The reason lies in the online nature of the problem: If every real item tries to replace an appropriate item in the model-packing it can happen that we have to open many new bins in the beginning of a phase. If the phase then suddenly stops we have possibly opened too many bins. Packing virtual items in a clever way using the auxiliary algorithm we are able to avoid this problem.

Let $A(P_i)$ denote the number of opened bins by A for items from phase P_i. There are two different reasons for opening a bin:

- The auxiliary algorithm A' opens a new bin to pack the newly generated virtual item;
- The real item cannot replace a virtual item.

The number of opened bins for virtual items by the auxiliary algorithm is denoted by $A'(P_i)$ and the number of real items that cannot replace a virtual item by $U(P_i)$. So we have $A(P_i) = A'(P_i) + U(P_i)$. In the analysis we will independently control $A'(P_i)$ as well as $U(P_i)$.

Let $t = \lceil m/2 \rceil + 1$. The analysis is split up into three parts:

- We call the phases $P_0, \dots, P_{t-1}$ the *first phases*. To estimate the number of opened bins in the first phases we apply a worst-case estimate.
- In case that the last phase P_m is very short, that is, $|P_m| < \sqrt{n}$ we also apply a worst-case estimate.
- The interesting part is the analysis of the number of opened bins in the remaining phases $P_t, \dots, P_{m-1}$ (and P_m if $|P_m| \geq \sqrt{n}$).

For each real item our algorithm opens at most two bins: One for the generated virtual item and one for the real item if it cannot replace a virtual item. Hence, the number of opened bins by A in the phases $P_0, \ldots, P_{t-1}$ (and the bin for A_1) is upper bounded by

$$1 + \sum_{i=0}^{t-1} 2|P_i| \leq 2 + 2\sum_{i=0}^{t-1} 2^i \leq 2 \cdot 2^t \leq 8 \cdot 2^{m/2} \leq 8\sqrt{n}.$$

Furthermore, if $|P_m| < \sqrt{n}$, then the number of opened bins for items in the last phase is upper bounded by $2\sqrt{n}$.

Now, we can proceed with the analysis of the remaining phases. We note that if we deal with a remaining phase P_i we have $|P_i^*| \geq |P_i| \geq \sqrt{n}$. The first important observation is the existence of an auxiliary algorithm A' that can pack the generated virtual items well. The following statement is the adaption of Theorem 2.1 in [72] from bin packing in the setup of sampling with replacement to cardinality-constrained bin packing in the setup of *sampling without replacement*:

▶ **Lemma 38.** *We consider cardinality-constrained bin packing with parameter $k \in \mathbb{N}$. Assume we are given a multiset $\mathcal{S} = \{a_1, \ldots, a_q\}$ of q items. Let $p \in \mathbb{N}$ with $p \leq q$. We generate a random instance I by drawing p items from $\mathcal{S}$ according to sampling without replacement. Then, there exists a (semi-)online algorithm A', which knows $\mathcal{S}$ in advance, but is unaware of p and I, and universal constants α, C, K such that we have*

$$\mathbb{P}\left[A'(I) \leq \frac{p}{q}\,\mathrm{OPT}(\mathcal{S}) + k \cdot K\sqrt{p}\log(p)^{3/4}\right] \geq 1 - k \cdot C\exp\left(-\alpha\log(p)^{3/2}\right).$$

The second important observation is that for most of the real items it is possible to replace a virtual item:

▶ **Lemma 39.** *There exist universal constants α, C, K such that we have*

$$\mathbb{P}\left[U(P_i) \leq K\sqrt{|P_i|}\log(|P_i|)^{3/4}\right] \geq 1 - C\exp\left(-\alpha\log(|P_i|)^{3/2}\right).$$

A proof of the two previous statements will be given in Section 3.2.1.2 and Section 3.2.1.3.

Combining Lemma 38 and Lemma 39 we obtain an estimate of the number of opened bins by A in phase P_i with high probability:

▶ **Lemma 40.** *There exist universal constants α, C, K such that the event*

$$\mathcal{G}_i := \left\{A(P_i) \leq \frac{|P_i|}{|\mathcal{I}|}\,\mathrm{OPT}(\mathcal{I}) + k \cdot K\sqrt{|P_i^*|}\log\left(|P_i^*|\right)^{3/4}\right\}$$

takes place with probability at least $1 - k \cdot C\exp\left(-\alpha\log\left(|P_i|\right)^{3/2}\right)$.

The proof of this lemma will be given in Section 3.2.1.3. Applying a union bound we see that the event $\mathcal{G}_t \cap \ldots \cap \mathcal{G}_m$ takes place with probability at least

$$\begin{aligned} 1 - \sum_{i=t}^{m} k \cdot C\exp\left(-\alpha\log\left(|P_i|\right)^{3/2}\right) &\geq 1 - \sum_{i=t}^{m} k \cdot C\exp\left(-\alpha\log\left(\sqrt{n}\right)^{3/2}\right) \\ &\geq 1 - k \cdot C\sum_{i=t}^{m}\exp\left(-\alpha\log(n)^{3/2}\right) \\ &\geq 1 - \log(n) \cdot k \cdot C\exp\left(-\alpha\log\left(n\right)^{3/2}\right) \end{aligned}$$

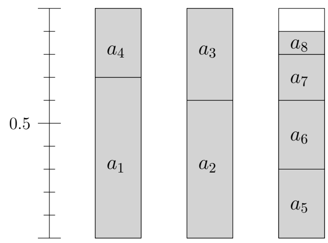

Figure 13 A model-packing $\mathcal{M}$ for the set of items $\{a_1, \ldots, a_8\}$ with $a_1 = 0.7$, $a_2 = 0.6$, $a_3 = 0.4$, $a_4 = a_5 = a_6 = 0.3$, $a_7 = 0.2$ and $a_8 = 0.1$. We see that $r(a_1) = r(a_2) = r(a_5) = 1$, $r(a_3) = r(a_4) = r(a_6) = 2$, $r(a_7) = 3$ and $r(a_8) = 4$.

$$\geq 1 - k \cdot C \exp\left(-\alpha \log(n)^{3/2}\right).$$

Here and in the future the used constants α, C and K may differ from line to line. But they are always universal, that is independent of the input.

Now assume that $\mathcal{G}_t \cap \ldots \cap \mathcal{G}_m$ takes place. Then, we have

$$\begin{aligned}
A(I^\sigma) &\leq 8\sqrt{n} + \sum_{i=t}^{m} \left(\frac{|P_i|}{|\mathcal{I}|} \operatorname{OPT}(\mathcal{I}) + k \cdot K \sqrt{|P_i^*| \log\left(|P_i^*|\right)^{3/4}} \right) \\
&\leq 8\sqrt{n} + \operatorname{OPT}(\mathcal{I}) + k \cdot K \sum_{i=t}^{m} \sqrt{2^i} \log\left(2^i\right)^{3/4} \\
&\leq 8\sqrt{n} + \operatorname{OPT}(\mathcal{I}) + k \cdot K \log(n)^{3/4} \sum_{i=t}^{m} 2^{i/2} \\
&\leq 8\sqrt{n} + \operatorname{OPT}(\mathcal{I}) + k \cdot K \log(n)^{3/4} \cdot 2^{m/2} \\
&\leq 8\sqrt{n} + \operatorname{OPT}(\mathcal{I}) + k \cdot K \sqrt{n} \log(n)^{3/4}.
\end{aligned}$$

If $\mathcal{G}_t \cap \ldots \cap \mathcal{G}_m$ does not take place, then we estimate $A(I^\sigma)$ by the number of opened bins in the worst case, that is $2n$. Using the inequality $\operatorname{OPT}(\mathcal{I}) \geq n/k$ we obtain

$$\begin{aligned}
\frac{\mathbb{E}\left[A(I^\sigma)\right]}{\operatorname{OPT}(\mathcal{I})} &\leq 1 \cdot \frac{\operatorname{OPT}(\mathcal{I}) + k \cdot K \sqrt{n} \log(n)^{3/4}}{\operatorname{OPT}(\mathcal{I})} + k \cdot C \exp\left(-\alpha \log(n)^{3/2}\right) \cdot \frac{2n}{\operatorname{OPT}(\mathcal{I})} \\
&\leq 1 + \frac{k \cdot K \sqrt{n} \log(n)^{3/4}}{n/k} + k \cdot C \exp\left(-\alpha \log(n)^{3/2}\right) \cdot \frac{2n}{n/k} \\
&\leq 1 + k^2 \cdot K \frac{\log(n)^{3/4}}{\sqrt{n}} + k^2 \cdot C \exp\left(-\alpha \log(n)^{3/2}\right).
\end{aligned}$$

This term tends to one as $\operatorname{OPT}(\mathcal{I})$, and so n, tends to infinity.

3.2.1.2 Description and Analysis of the Auxiliary Algorithm A'

Here we will prove:

▶ **Lemma 38.** *We consider cardinality-constrained bin packing with parameter $k \in \mathbb{N}$. Assume we are given a multiset $\mathcal{S} = \{a_1, \ldots, a_q\}$ of q items. Let $p \in \mathbb{N}$ with $p \leq q$. We generate a random instance I by drawing p items from $\mathcal{S}$ according to sampling without replacement. Then, there exists a (semi-)online algorithm A', which knows $\mathcal{S}$ in advance, but is unaware of p and I, and universal constants α, C, K such that we have*

$$\mathbb{P}\left[A'(I) \leq \frac{p}{q} \operatorname{OPT}(\mathcal{S}) + k \cdot K \sqrt{p} \log(p)^{3/4}\right] \geq 1 - k \cdot C \exp\left(-\alpha \log(p)^{3/2}\right).$$

The auxiliary algorithm A' proceeds as in [72]: Let $\mathcal{S} = \{a_1, \ldots, a_q\}$. Without loss of generality we can assume that all items have different sizes, otherwise we distinguish items of the same size by their index. Before starting to pack the items, we compute an optimal packing $\mathcal{M}(\mathcal{S})$ of the items in $\mathcal{S}$, the so-called *model-packing*. An example is given in Figure 13. Then, with respect to a fixed model-packing, we establish a *rank*-function $r : \{a_1, \ldots, a_q\} \to [k]$, where $r(a_i)$ is equal to ℓ, if a_i is the ℓ-th largest item in its bin in the model-packing.

Now, let $I = (V_1, \ldots, V_p)$ denote the random instance drawn by sampling p items from $\mathcal{S}$ according to sampling without replacement. A' also uses the concept of virtual items that serve as placeholders. In order to distinguish the virtual items introduced by A' from the virtual items introduced by A we call these items *second-order virtual items.* Assume now that a new item V_j arrives that has to be packed by A'. We know that $V_j = a_m$ for a unique $m \in \{1, \ldots, q\}$. We distinguish two cases for the behavior of A':

- At first we assume that $r(a_m) = 1$. Let $b_1 > b_2 > \ldots > b_\ell$ denote the items in the bin containing a_m in the model-packing (so we have $b_1 = a_m$). We assume that $r(b_2) = 2$, $r(b_3) = 3$, and so on. Then, A' opens a new bin and puts V_j into it. Furthermore, we add second-order virtual items $W_2, \ldots, W_\ell$ with $W_2 = b_2, \ldots, W_\ell = b_\ell$ to this bin.
- Now, assume that we have $1 < r(a_m) =: r$. Then, we try to replace the smallest second-order virtual item W in an opened bin with $r(W) = r$ and $W \geq V_j$ by V_j. If such an item W does not exist, we put V_j into a new bin.

Let O_u with $u \in [k]$ denote the number of opened bins by A' for items with rank u. Then, we have $A'(I) = \sum_{u=1}^{k} O_u$.

The analysis of $A'(I)$ is similar to the one given in [72], we only have to use the adjusted matching-statements for the case of sampling without replacement. In fact, the analysis is even a bit less involved, since the constraint that at most k items are allowed in a bin makes it possible to use a simple union bound.

There are $\mathrm{OPT}(\mathcal{S})$ many items with rank one among the items in $\mathcal{S}$. So the expected number of opened bins by items of rank one after drawing p many items with respect to sampling without replacement is $\frac{p}{q} \cdot \mathrm{OPT}(\mathcal{S})$. Applying Hoeffding's inequality (see Proposition 12) we obtain

$$\mathbb{P}\left[O_1 \leq \frac{p}{q}\,\mathrm{OPT}(\mathcal{S}) + \sqrt{p}\log(p)^{3/4}\right] \geq 1 - \exp\left(-2\log(p)^{3/2}\right).$$

It remains to bound the number of opened bins by items V_i of rank u, with $u \in \{2, \ldots, k\}$, from above. In order to this, we consider the case of a fixed u. Then, we have a stochastic process that consists of two possible events:

- An item of rank 1, whose corresponding bin in the model-packing contains at least u items, is drawn;
- An item of rank u is drawn.

If an item of rank 1 is drawn, a new second-order virtual item is generated, which offers space for new items of rank u. If an item of rank u is drawn, we try to replace an existing second-order virtual item of rank u. Without loss of generality we can assume that all items of rank u in the model packing have different sizes, otherwise we only have more possibilities to replace second-order virtual items. The number of opened bins due to rank u items corresponds exactly to the number of rank u items that cannot replace a second-order virtual item. But this question is equivalent to the introduced matching variant (M2), where we treat the virtual items of rank u as plus-points, and the second-order virtual items generated by items of rank 1 as minus-points. Hence, we have $n := \#(\text{items of rank } u \text{ in the model-packing})$ and a random instance size L_u.

We distinguish two cases: At first we assume that $L_u \leq p^{1/3}$. Then, we also have $O_u \leq p^{1/3}$. So assume that $L_u > p^{1/3}$. Then, it follows from Lemma 25 that we have

$$O_u \leq K \cdot \sqrt{L_u} \log(L_u)^{3/4} \leq K \cdot \sqrt{p} \log(p)^{3/4}$$

with probability at least

$$1 - C \exp\left(-\alpha \log(L_u)^{3/2}\right) \geq 1 - C \exp\left(-\alpha \log(p^{1/3})^{3/2}\right) \geq 1 - C \exp\left(-\alpha \log(p)^{3/2}\right).$$

It follows that we have $O_u \leq p^{1/3} + K \cdot \sqrt{p} \log(p)^{3/4} \leq K \cdot \sqrt{p} \log(p)^{3/4}$ with probability at least $1 - C \exp\left(-\alpha \log(p)^{3/2}\right)$.

Using a union bound on the ranks $2, \dots, k$, we see that

$$A'(I) \leq \frac{p}{q} \operatorname{OPT}(\mathcal{S}) + k \cdot K \sqrt{p} \log(p)^{3/4}$$

with probability at least

$$1 - \exp\left(-2 \log(p)^{3/2}\right) - (k-1) \cdot C \exp\left(-\alpha \log(p)^{3/2}\right) \geq 1 - k \cdot C \exp\left(-\alpha \log(p)^{3/2}\right).$$

3.2.1.3 Proof of Lemma 39 and 40

Here we will prove:

▶ **Lemma 39.** *There exist universal constants α, C, K such that we have*

$$\mathbb{P}\left[U(P_i) \leq K \sqrt{|P_i|} \log(|P_i|)^{3/4}\right] \geq 1 - C \exp\left(-\alpha \log(|P_i|)^{3/2}\right).$$

and

▶ **Lemma 40.** *There exist universal constants α, C, K such that the event*

$$\mathcal{G}_i := \left\{ A(P_i) \leq \frac{|P_i|}{|\mathcal{I}|} \operatorname{OPT}(\mathcal{I}) + k \cdot K \sqrt{|P_i^*|} \log\left(|P_i^*|\right)^{3/4} \right\}$$

takes place with probability at least $1 - k \cdot C \exp\left(-\alpha \log\left(|P_i|\right)^{3/2}\right)$.

To prove Lemma 39, we observe that we can assume without loss of generality that all items have different sizes, otherwise we only take advantage. Then, we see that the question of how many real items cannot replace a virtual item is equivalent to matching-variant (M1), where the virtual items are represented by the minus-points and the real items are represented by the plus-points.

To prove Lemma 40, we apply Lemma 38 one time with $\mathcal{S} = P_i^*$ and $p = |P_i|$ and one time with $\mathcal{S} = \mathcal{I}$ and $p = |P_i^*|$. Then, we obtain

$$\mathbb{P}\left[A'(P_i) \leq \frac{|P_i|}{|P_i^*|} \cdot \operatorname{OPT}(P_i^*) + k \cdot K \sqrt{|P_i|} \log(|P_i|)^{3/4}\right]$$

$$\geq 1 - k \cdot C \exp\left(-\alpha \log(|P_i|)^{3/2}\right)$$

and

$$\mathbb{P}\left[\mathrm{OPT}(P_i^*) \leq \frac{|P_i^*|}{|\mathcal{I}|} \cdot \mathrm{OPT}(\mathcal{I}) + k \cdot K\sqrt{|P_i^*|}\log(|P_i^*|)^{3/4}\right]$$
$$\geq 1 - k \cdot C \exp\left(-\alpha \log(|P_i^*|)^{3/2}\right) \geq 1 - k \cdot C \exp\left(-\alpha \log(|P_i|)^{3/2}\right).$$

Combining these two estimates with Lemma 39 we obtain the statement of Lemma 40.

3.2.2 Existence of a Nearly Optimal Algorithm for Classical Bin Packing with respect to the Random-order Ratio

Here, we will prove:

▶ **Theorem 34.** *For online classical bin packing there exists for every ϵ greater than zero a randomized algorithm A_ϵ with $\mathrm{RR}(A_\epsilon) \leq 1 + \epsilon$.*

3.2.2.1 The Setting and the Algorithm

The setting is similar to the case of cardinality-constrained bin packing: Again, we are given a multiset $\mathcal{I} = \{a_1, \ldots, a_n\}$ of items. Let $\gamma > 0$ be a parameter that we will choose, depending on ϵ, later on. Items with sizes greater than or equal to γ are called *large*, and otherwise *small*. We set $\mathcal{I}_{\geq\gamma} = \{a : a \in \mathcal{I}, a \geq \gamma\}$ and $\mathcal{I}_{<\gamma} = \{a : a \in \mathcal{I}, a < \gamma\}$. The random instance we are dealing with is then given by $I^\sigma = (A_1, \ldots, A_n)$ with $A_i = a_{\sigma(i)}$, where σ is a random permutation (that is, random with respect to the uniform distribution on $[n]$).

We observe that the number of large items in a bin is bounded from above by $\frac{1}{\gamma}$. Hence, packing large items is similar to cardinality-constrained bin packing with parameter $k = \lfloor\gamma^{-1}\rfloor$ (apart from the fact that the number of large items in each phase is random). To pack the small items we observe that if we open bins with the auxiliary algorithms there possibly arises some empty space, which is not occupied by virtual items: We want to use this space (and empty space from previous phases) by packing the small items using First-Fit into these gaps.

We partition the input in phases $P_0, \ldots, P_m$ as in the case of cardinality-constrained bin packing: The i-th phase covers the items $(A_{2^i+1}, \ldots, A_{\min\{2^{i+1},n\}})$. Moreover, $P_i^* = \{A_1, \ldots, A_{2^i}\}$ denotes the multiset of all items drawn before phase P_i starts. Let L_i denote the number of large items drawn in phase P_i and L_i^* the number of large items in P_i^*.

The first item A_1 will be packed separately. Now we want to describe how the algorithm packs the items in phase P_i. At the beginning of phase P_i we delete all virtual items generated in phase P_{i-1} that are remaining. Bins that received only virtual items in the past and are empty after deleting them, still count as opened and can be used in the following for accommodating small items. Then, we compute an optimal packing of all *large* items in P_i^*, the model-packing $\mathcal{M}(P_i^*)$. Furthermore, let $\mathcal{Z}_i$ denote the multiset of all large items drawn in P_i^*. Let A_j denote the arriving item that has to be packed. We distinguish whether A_j is a large or a small item:

- At first we assume that A_j is a large item. In this case, we want to proceed as in cardinality-constrained bin packing. If $\mathcal{Z}_i \neq \emptyset$, then we generate a virtual item V_j as follows: We draw a random item R from $\mathcal{Z}_i$ and set $V_j := R$. Afterwards, we set $\mathcal{Z}_i := \mathcal{Z}_i \backslash \{R\}$. Then, we pack V_j using the auxiliary algorithm described in the section about cardinality-constrained bin packing. Finally, we try to replace the smallest virtual item V with

$V \geq A_j$, which is not already replaced by a real item. If no such virtual item exists, then we open a new bin for A_j and the remaining space in this bin is fully designated for small items. If $\mathcal{Z}_i = \emptyset$, then we open a new bin for A_j and the remaining space in this bin will be also fully designated for small items.
- Now we assume that A_j is a small item. We observe that there can be empty space in all bins that are opened *before* phase P_i. Since those bins will not receive more large items, we can try to use their remaining space for small items. Moreover, we observe that in the model-packing, the bins are not necessarily filled up to their capacity. For example, the bin on the right hand side in the model-packing shown in Figure 13 is only filled up to 0.9. The remaining space is now intended to be filled up with small items. That is, every time the auxiliary algorithm packs a large item of rank 1, we open some empty space (which can be also zero) designated for small items. Then, we try to pack the small items using First-Fit into the designated space. If that is not possible, we open a new bin for the currently processed small item and the space in this bin is fully designated for small items.

We want to emphasize once more that while large items are still packed independently for each phase this is not true for small items.

3.2.2.2 High-Level Proof

Here, we give a high-level proof of the statement. The proofs of the arising lemmata are postponed to the following sections.

We will consider the number of opened bins for large items and the number of opened bins for small items separately. $A^{\ell}(P_i)$ denotes the number of opened bins by A due to *large* items in phase P_i and $A^s(P_i)$ the number of opened bins due to *small* items. Before starting to analyze the number of opened bins for large items, we will show that we can assume that $|\mathcal{I}_{\geq\gamma}|$ grows linearly in $\mathrm{OPT}(\mathcal{I})$. For a random instance I^σ let $L(I^\sigma)$ denote the number of opened bins by A for large items. We make the following observations:

- $L(I^\sigma)$ is bounded from above by $2|\mathcal{I}_{\geq\gamma}|$. This is, because for each large item at most two bins are opened: One may be opened by the auxiliary algorithm to pack the generated virtual item, and one since we are not able to replace a virtual item by a real one. Furthermore, we have $L(I^\sigma) \geq \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) \geq S(\mathcal{I}_{\geq\gamma}) \geq \gamma|\mathcal{I}_{\geq\gamma}|$, where $S(\mathcal{I}_{\geq\gamma})$ denotes the total size of all items in $\mathcal{I}_{\geq\gamma}$.
- Since we pack the small items using First-Fit, the number of opened bins caused by small items is bounded from above by $(1-\gamma)^{-1}S(\mathcal{I}_{<\gamma}) + 1$.

Now assume that $|\mathcal{I}_{\geq\gamma}| \leq \gamma\,\mathrm{OPT}(\mathcal{I})$. Then, (using the inequality $(1-\gamma)^{-1} \leq 1 + 2\gamma$ for $\gamma \leq 1/2$) we have

$$\frac{A(I^\sigma)}{\mathrm{OPT}(\mathcal{I})} \leq \frac{2|\mathcal{I}_{\geq\gamma}| + (1-\gamma)^{-1}S(\mathcal{I}_{<\gamma}) + 1}{\mathrm{OPT}(\mathcal{I})} \leq \frac{2\gamma\,\mathrm{OPT}(\mathcal{I}) + (1+2\gamma)\,\mathrm{OPT}(\mathcal{I}) + 1}{\mathrm{OPT}(\mathcal{I})} \leq 1 + 4\gamma + \frac{1}{\mathrm{OPT}(\mathcal{I})}.$$

So in this case, if we choose γ sufficiently small depending on ϵ, our algorithm would be asymptotically $(1+\epsilon)$-competitive. Hence, from now on we will make the following assumption:

▶ **Assumption 1.** *We have* $|\mathcal{I}_{\geq\gamma}| > \gamma\,\mathrm{OPT}(\mathcal{I})$.

At first, we consider the number of opened bins for large items. This time we set the period of first phases larger and set $t = \lceil \log_2(\gamma^2 n) \rceil$.

▶ **Lemma 41.** *There exist universal positive constants α_γ and C such that the event*

$$\mathcal{G}^\ell_{start} := \left\{ \sum_{i=0}^{t-1} A^\ell(P_i) \le 6\gamma \, \mathrm{OPT}(\mathcal{I}) \right\}$$

takes place with probability at least $1 - C \exp\left(-\alpha_\gamma \, \mathrm{OPT}(\mathcal{I})\right)$.

Furthermore, if $|P_m| < \gamma^2 n$, then also exist universal positive constants α_γ and C such that the event

$$\mathcal{G}^\ell_{end} := \left\{ A^\ell(P_m) \le 4\gamma \, \mathrm{OPT}(\mathcal{I}) \right\}$$

takes place with probability at least $1 - C \exp\left(-\alpha_\gamma \, \mathrm{OPT}(\mathcal{I})\right)$.

We will now assume that the length of the last phase is at least $\gamma^2 n$. Then, we obtain for the remaining phases:

▶ **Lemma 42.** *There exist universal positive constants α_γ, C_γ and K_γ such that*

$$\mathcal{G}^\ell_{remaining} = \left\{ \sum_{i=t}^{m} A^\ell(P_i) \le \mathrm{OPT}(\mathcal{I}_{\ge\gamma}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4} \right\}$$

takes place with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$.

Now, we consider the opened bins for small items. Let $S(\mathcal{I}_{<\gamma})$, $S(\mathcal{I}_{\ge\gamma})$, and $S(\mathcal{I})$ denote the total size of all items in the corresponding sets. At first we show that we can assume that $S(\mathcal{I}_{<\gamma}) > \gamma \, \mathrm{OPT}(\mathcal{I})$: Assume that $S(\mathcal{I}_{<\gamma}) \le \gamma \, \mathrm{OPT}(\mathcal{I})$. Then, we have (for $\gamma \le 1/2$)

$$\begin{aligned}
\frac{A(I^\sigma)}{\mathrm{OPT}(\mathcal{I})} &\le \frac{L(I^\sigma) + (1-\gamma)^{-1} S(\mathcal{I}_{<\gamma}) + 1}{\mathrm{OPT}(\mathcal{I})} \\
&\le \frac{L(I^\sigma) + (1+2\gamma) \cdot \gamma \, \mathrm{OPT}(\mathcal{I}) + 1}{\mathrm{OPT}(\mathcal{I})} \\
&\le \frac{L(I^\sigma)}{\mathrm{OPT}(\mathcal{I})} + 3\gamma + \frac{1}{\mathrm{OPT}(\mathcal{I})}.
\end{aligned}$$

We still assume that $|P_m| \ge \gamma^2 n$, otherwise, we can proceed in a similar way. Then, using a union bound it follows from Lemma 41 and 42 that $\mathcal{G}^\ell_{\text{start}} \cap \mathcal{G}^\ell_{\text{remaining}}$ takes place with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$. In case that this event does not take place we use the worst-case estimate $L(I^\sigma) \le 2|\mathcal{I}_{\ge\gamma}| \le \frac{2}{\gamma} \mathrm{OPT}(\mathcal{I}_{\ge\gamma})$. Then, we obtain

$$\begin{aligned}
\frac{\mathbb{E}\left[A(I^\sigma)\right]}{\mathrm{OPT}(\mathcal{I})} &\le \frac{\mathrm{OPT}(\mathcal{I}_{\ge\gamma}) + 6\gamma \, \mathrm{OPT}(\mathcal{I}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I})^{3/4}}{\mathrm{OPT}(\mathcal{I})} \\
&\qquad + \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right) \cdot \frac{\frac{2}{\gamma}\mathrm{OPT}(\mathcal{I})}{\mathrm{OPT}(\mathcal{I})} + 3\gamma + \frac{1}{\mathrm{OPT}(\mathcal{I})} \\
&\le 1 + 9\gamma + K_\gamma \frac{\log(\mathrm{OPT}(\mathcal{I}))^{3/4}}{\sqrt{\mathrm{OPT}(\mathcal{I})}} + \frac{2}{\gamma} \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right) + \frac{1}{\mathrm{OPT}(\mathcal{I})}.
\end{aligned}$$

The last three terms tend to zero as $\mathrm{OPT}(\mathcal{I})$ tends to infinity. So, choosing γ appropriately depending on ϵ, it follows that our algorithm would be asymptotically $(1+\epsilon)$-competitive in the random-order model. Hence, our second assumption is as follows:

▶ **Assumption 2.** *It holds* $S(\mathcal{I}_{<\gamma}) > \gamma \operatorname{OPT}(\mathcal{I})$.

We begin by estimating the number of opened bins for small items in the first phases (and in the last phase if the last phase is short). This time it does not suffice to bound the *number* of small items in the first phases but we have to estimate their *total* size.

▶ **Lemma 43.** *There exist universal positive constants* C *and* α_γ *such that the event*

$$\mathcal{G}^s_{start} := \left\{ \sum_{i=0}^{t-1} A^s(P_i) \leq 9\gamma \operatorname{OPT}(\mathcal{I}) + 1 \right\}$$

takes place with probability at least $1 - C \exp(-\alpha_\gamma \operatorname{OPT}(\mathcal{I}))$.

Furthermore, if $|P_m| < \gamma^2 n$*, then also exist universal positive constants* α_γ *and* C *such that the event*

$$\mathcal{G}^s_{end} := \{A^s(P_m) \leq 6\gamma \operatorname{OPT}(\mathcal{I}) + 1\}$$

takes place with probability at least $1 - C \exp(-\alpha_\gamma \operatorname{OPT}(\mathcal{I}))$.

Now we analyze the remaining phases: Let $I^\sigma_{<\gamma}$ denote an arbitrary permutation of the small items. Assume we generate a packing for $\mathcal{I}$ as follows: At first we pack the items from $\mathcal{I}_{\geq\gamma}$ in an optimal way, that is, in a way such that we need $\operatorname{OPT}(\mathcal{I}_{\geq\gamma})$ many bins. Then, we add the items from $I^\sigma_{<\gamma}$ using First-Fit to the existent packing. Let $\operatorname{FF}(I^\sigma_{<\gamma})$ denote the number of opened bins by First-Fit.

It follows as in the proof of the existence of a PTAS for bin packing that we have (for $\gamma \leq 1/2$) for arbitrary $I^\sigma_{<\gamma}$

$$\operatorname{OPT}(\mathcal{I}_{\geq\gamma}) + \operatorname{FF}(I^\sigma_{<\gamma}) \leq \left\lceil \frac{1}{1-\gamma} \operatorname{OPT}(\mathcal{I}) \right\rceil + 1 \leq (1+2\gamma) \operatorname{OPT}(\mathcal{I}) + 2. \tag{6}$$

Furthermore, we observe that it holds

$$\operatorname{FF}(I^\sigma_{<\gamma}) \geq [S(\mathcal{I}_{<\gamma}) - (\operatorname{OPT}(\mathcal{I}_{\geq\gamma}) - S(\mathcal{I}_{\geq\gamma}))]^+ = [S(\mathcal{I}) - \operatorname{OPT}(\mathcal{I}_{\geq\gamma})]^+, \tag{7}$$

where $[x]^+ := \max\{x, 0\}$. This follows, because $\operatorname{OPT}(\mathcal{I}_{\geq\gamma}) - S(\mathcal{I}_{\geq\gamma})$ is equal to the remaining space that can receive small items.

The following lemma is the central part of our proof of Theorem 34:

▶ **Lemma 44.** *There exist universal positive constants* α_γ, C_γ *and* K_γ *such that the event* $\mathcal{G}^s_{remaining}$ *defined as*

$$\left\{ \sum_{i=t}^{m} A^s(P_i) \leq [S(\mathcal{I}) - \operatorname{OPT}(\mathcal{I}_{\geq\gamma})]^+ + 7\gamma \operatorname{OPT}(\mathcal{I}) + K_\gamma \sqrt{\operatorname{OPT}(\mathcal{I})} \log(\operatorname{OPT}(\mathcal{I}))^{3/4} \right\}$$

takes place with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\operatorname{OPT}(\mathcal{I}))^{3/2}\right)$.

Again, we assume that $|P_m| \geq \gamma^2 n$. Then, using the previous lemmata we see that there exist universal positive constants α_γ, C_γ and K_γ such that the event $\mathcal{G}^\ell_{\text{start}} \cap \mathcal{G}^\ell_{\text{remaining}} \cap \mathcal{G}^s_{\text{start}} \cap \mathcal{G}^s_{\text{remaining}}$ takes place with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\operatorname{OPT}(\mathcal{I}))^{3/2}\right)$. In case that this event does not take place we use the worst-case estimate (for $\gamma \leq 1/2$)

$$\begin{aligned} A(I^\sigma) &\leq L(I^\sigma) + (1-\gamma)^{-1} S(\mathcal{I}_{<\gamma}) + 1 \\ &\leq L(I^\sigma) + (1+2\gamma) S(\mathcal{I}_{<\gamma}) + 1 \end{aligned}$$

$$\leq 2|\mathcal{I}_{\geq\gamma}| + (1+2\gamma)S(\mathcal{I}) + 1$$
$$\leq \frac{2}{\gamma}\,\mathrm{OPT}(\mathcal{I}) + (1+2\gamma)\,\mathrm{OPT}(\mathcal{I}) + 1$$
$$\leq \left(\frac{2}{\gamma} + 2 + 2\gamma\right)\mathrm{OPT}(\mathcal{I}).$$

Then, using (6) and (7) we obtain

$$\frac{\mathbb{E}\left[A(I^\sigma)\right]}{\mathrm{OPT}(\mathcal{I})} \leq \frac{\mathrm{OPT}(\mathcal{I}_{\geq\gamma}) + [S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})]^+ + 22\gamma\,\mathrm{OPT}(\mathcal{I})}{\mathrm{OPT}(\mathcal{I})}$$
$$+ \frac{K_\gamma\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}}{\mathrm{OPT}(\mathcal{I})}$$
$$+ \frac{\exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)\cdot\left(\frac{2}{\gamma}+2+2\gamma\right)\mathrm{OPT}(\mathcal{I})}{\mathrm{OPT}(\mathcal{I})}$$
$$\leq 1 + 24\gamma + K_\gamma\frac{\log(\mathrm{OPT}(\mathcal{I}))^{3/4}}{\sqrt{\mathrm{OPT}(\mathcal{I})}} + K_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right).$$

The last two terms tend to zero as $\mathrm{OPT}(\mathcal{I})$ tends to infinity. Thus, choosing γ sufficiently small our algorithm is asymptotically $(1+\epsilon)$-competitive in the random-order model.

3.2.2.3 Useful Inequalities

The following inequalities will be used extensively in the following:

- $\mathrm{OPT}(\mathcal{I}_{\geq\gamma}) \leq \mathrm{OPT}(\mathcal{I})$;
- $\mathrm{OPT}(\mathcal{I})/2 < S(\mathcal{I}) \leq \lceil S(\mathcal{I})\rceil \leq \mathrm{OPT}(\mathcal{I})$;
- $S(\mathcal{I}_{\geq\gamma}) \leq S(\mathcal{I})$ and $S(\mathcal{I}_{<\gamma}) \leq S(\mathcal{I})$;
- $|\mathcal{I}_{\geq\gamma}| \leq 1/\gamma\,\mathrm{OPT}(\mathcal{I}_{\geq\gamma}) \leq 1/\gamma\,\mathrm{OPT}(\mathcal{I})$.

The inequality $S(\mathcal{I}) > \mathrm{OPT}(\mathcal{I})/2$ can be shown in the following way: Let $\ell_1,\ldots,\ell_{\mathrm{OPT}(\mathcal{I})}$ denote the bin levels of a fixed optimal packing. Because of the optimality we have $\ell_i+\ell_j > 1$ for $i,j \in [\mathrm{OPT}(\mathcal{I})]$ with $i \neq j$. If $\mathrm{OPT}(\mathcal{I})$ is even, the statement follows immediately. If $\mathrm{OPT}(\mathcal{I})$ is odd, then we have $S(\mathcal{I}) = \ell_1+\ldots+\ell_{\mathrm{OPT}(\mathcal{I})-1}+\ell_{\mathrm{OPT}(\mathcal{I})} \geq (\mathrm{OPT}(\mathcal{I})-1)/2+\ell_{\mathrm{OPT}(\mathcal{I})}$ and $S(\mathcal{I}) \geq \ell_1 + (\mathrm{OPT}(\mathcal{I})-1)/2$. Therefore, we have $2S(\mathcal{I}) \geq \mathrm{OPT}(\mathcal{I}) - 1 + \ell_1 + \ell_{\mathrm{OPT}(\mathcal{I})} > \mathrm{OPT}(\mathcal{I})$.

Moreover, we will use the following estimates:

- For $\gamma \leq 1/2$ we have $(1-\gamma)^{-1} \leq 1+2\gamma$.
- Let α_γ and β_γ be two positive constants. Then, there exists another positive constant $\tilde{\alpha}_\gamma$ such that we have for ℓ large enough the inequality $\exp\left(-\alpha_\gamma\log(\beta_\gamma\cdot\ell)^{3/2}\right) \leq \exp\left(-\tilde{\alpha}_\gamma\log(\ell)^{3/2}\right)$.
- For a fixed $\alpha > 1$ we have for all ℓ large enough $\log(\alpha\ell) \leq \alpha\log(\ell)$.

3.2.2.4 Proof of Lemma 41

Here we will prove:

▶ **Lemma 41.** *There exist universal positive constants α_γ and C such that the event*

$$\mathcal{G}^\ell_{start} := \left\{\sum_{i=0}^{t-1} A^\ell(P_i) \leq 6\gamma\,\mathrm{OPT}(\mathcal{I})\right\}$$

takes place with probability at least $1 - C\exp(-\alpha_\gamma \operatorname{OPT}(\mathcal{I}))$.

Furthermore, if $|P_m| < \gamma^2 n$*, then also exist universal positive constants* α_γ *and* C *such that the event*

$$\mathcal{G}^\ell_{end} := \left\{A^\ell(P_m) \le 4\gamma \operatorname{OPT}(\mathcal{I})\right\}$$

takes place with probability at least $1 - C\exp(-\alpha_\gamma \operatorname{OPT}(\mathcal{I}))$.

Proof of Lemma 41. We set $b := |\mathcal{I}_{\ge\gamma}|$. For $i \in \{0,\dots,m\}$ let L_i denote the number of large items in phase P_i and L_i^* the number of large items in P_i^*. We have chosen t in a way that P_t is the first phase such that there are already at least $\gamma^2 n$ many items drawn. Furthermore, we have $|P_t^*| = 2^{\lceil \log_2(\gamma^2 n)\rceil} \le 2\gamma^2 n$, and therefore $\mathbb{E}[L_t^*] = |P_t^*| \cdot \frac{b}{n} \le 2\gamma^2 b$.

Let X_i be equal to one, if the i-th drawn item is large, and zero otherwise. Then, we have $L_t^* = \sum_{j=1}^{|P_t^*|} X_j$. Moreover, it is $\mathbb{E}\left[X_j^2\right] \le \mathbb{E}[X_j] = {}^b/_n$. Applying Bernstein's inequality (see Proposition 13) we obtain

$$\begin{aligned}\mathbb{P}\left[L_t^* \ge 3\gamma^2 b\right] &\le \mathbb{P}\left[L_t^* \ge \mathbb{E}[L_t^*] + \gamma^2 b\right]\\ &\le \exp\left(-\frac{[\gamma^2 b]^2}{2\cdot 2^{\lceil \log_2(\gamma^2 n)\rceil}\cdot \frac{b}{n} + \frac{2}{3}\cdot\gamma^2 b}\right)\\ &\le \exp\left(-\frac{\gamma^4 b^2}{4\gamma^2 n\cdot\frac{b}{n} + \gamma^2 b}\right)\\ &\le \exp\left(-\frac{1}{5}\cdot\gamma^2 b\right)\\ &\le \exp\left(-\frac{1}{5}\cdot\gamma^3 \operatorname{OPT}(\mathcal{I})\right).\end{aligned}$$

Here, the last inequality follows from Assumption 1. Since the algorithm opens at most two bins per large item we have an upper bound for the number of opened bins of

$$2\cdot 3\gamma^2 b \le 6\gamma \operatorname{OPT}(\mathcal{I}).$$

This shows the first part of the lemma.

To analyze the last phase P_m in case of $|P_m| < \gamma^2 n$ we proceed as in the analysis of the first phases: It holds $\mathbb{E}[L_m] < \gamma^2 n \cdot \frac{b}{n} = \gamma^2 b$. Applying Bernstein's inequality we see that

$$\begin{aligned}\mathbb{P}\left[L_m \ge 2\gamma^2 b\right] \le \mathbb{P}\left[L_m \ge \mathbb{E}[L_m] + \gamma^2 b\right] &\le \exp\left(-\frac{[\gamma^2 b]^2}{2\cdot\gamma^2 n\cdot\frac{b}{n} + \frac{2}{3}\gamma^2 b}\right)\\ &\le \exp\left(-\frac{1}{3}\gamma^2 b\right) \le \exp\left(-\frac{1}{3}\gamma^3 \operatorname{OPT}(\mathcal{I})\right).\end{aligned}$$

Hence, if we assume that for each large item in the last phase two bins are opened, then with probability at least $1 - \exp\left(-\frac{1}{3}\gamma^3 \operatorname{OPT}(\mathcal{I})\right)$ we open at most $4\gamma \operatorname{OPT}(\mathcal{I})$ many bins for large items in the last phase. ◀

3.2.2.5 Proof of Lemma 42

Here we will prove

▶ **Lemma 42.** *There exist universal positive constants* α_γ, C_γ *and* K_γ *such that*

$$\mathcal{G}^\ell_{remaining} = \left\{\sum_{i=t}^{m} A^\ell(P_i) \le \operatorname{OPT}(\mathcal{I}_{\ge\gamma}) + K_\gamma\sqrt{\operatorname{OPT}(\mathcal{I})}\log(\operatorname{OPT}(\mathcal{I}))^{3/4}\right\}$$

takes place with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$.

Proof of Lemma 42. Now we consider the phases $P_t, \ldots, P_{m-1}$ and P_m in the case that $|P_m| \geq \gamma^2 n$. Let $i \in \{t, \ldots, m\}$. We see that there are three reasons to open a new bin because of a large item:

- If $L_i > L_i^*$, then we are not able to generate virtual items for the last $L_i - L_i^*$ many arriving large items and therefore pack each of them in a new bin;
- The auxiliary algorithm opens a new bin;
- We cannot replace a virtual item by the arriving real item.

Notice that since $t = \lceil \log_2(\gamma^2 n) \rceil$ we have

$$m - t + 1 \leq \log_2(n/2) + 2 - \log_2(\gamma^2 n) = 1 + \log_2\left(\frac{1}{\gamma^2}\right),$$

that is, the number of phases we consider now is upper bounded by a constant C_γ, which depends only on γ. This allows us to give the same bounds for all the phases and applying a union bound afterwards.

Again, we set $b := |\mathcal{I}_{\geq\gamma}|$. Let

$$\mathcal{E}_L = \bigcap_{i=t}^{m} \left(\left\{ |L_i - \mathbb{E}[L_i]| \leq \sqrt{b} \log(b)^{3/4} \right\} \cap \left\{ |L_i^* - \mathbb{E}[L_i^*]| \leq \sqrt{b} \log(b)^{3/4} \right\} \right)$$

denote the event that the number of large items in $P_t, \ldots, P_m$ and $P_t^*, \ldots, P_m^*$ is concentrated around its mean. At first we assume that $\mathcal{E}_L$ takes place, and analyze the number of opened bins for large items. Later on, we will show that $\mathcal{E}_L$ takes place with sufficient probability.

If $\mathcal{E}_L$ takes place, then for $i \in \{t, \ldots, m\}$ it holds true that $L_i - L_i^* \leq 2\sqrt{b} \log(b)^{3/4}$. That is, in this case the number of opened bins since we are not able to generate virtual items is upper bounded by $2\sqrt{b}\log(b)^{3/4}$ and in all phases together by

$$C_\gamma \sqrt{b} \log(b)^{3/4} \leq C_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}$$

for $\mathrm{OPT}(\mathcal{I})$ large enough.

Now, we study the number of opened bins because a real item cannot replace a virtual item: We denote this random variable again as $U(P_i)$. We see that here we have in phase P_i an instance of (M1) with $k = \min\{L_i, L_i^*\}$. Moreover, since we assume that $b \geq \gamma\,\mathrm{OPT}(\mathcal{I})$ and that $\mathrm{OPT}(\mathcal{I})$ is very large, we can assume that $\sqrt{b}\log(b)^{3/4} \leq \frac{1}{2} \cdot \frac{|P_i|}{n} \cdot b$. Then, it follows from Lemma 25 that in phase P_i we have

$$\begin{aligned}
\mathbb{P}\left[U(P_i) \leq K\sqrt{b}\log(b)^{3/4}\right] &\geq \mathbb{P}\left[U(P_i) \leq K\sqrt{\min\{L_i, L_i^*\}} \log(\min\{L_i, L_i^*\})^{3/4}\right] \\
&\geq 1 - C\exp\left(-\alpha \log\left(\min\{L_i, L_i^*\}\right)^{3/2}\right) \\
&\geq 1 - C\exp\left(-\alpha \log\left(\min\left\{\frac{1}{2} \cdot \frac{|P_i|}{n} \cdot b, \frac{1}{2} \cdot \frac{|P_i^*|}{n} \cdot b\right\}\right)^{3/2}\right) \\
&\geq 1 - C\exp\left(-\alpha \log\left(\frac{1}{2}\gamma^2 b\right)^{3/2}\right) \\
&\geq 1 - C\exp\left(-\alpha_\gamma \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/2}\right).
\end{aligned}$$

So, applying a union bound we see that

$$\mathbb{P}\left[\sum_{i=t}^{m} U(P_i) \leq K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right] \geq 1 - C_\gamma \exp\left(-\alpha_\gamma \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/2}\right).$$

Finally, we control the number of opened bins by the auxiliary algorithm. In order to do this, we apply two times Lemma 38: At first it follows that we have for $i \in \{t, \ldots, m\}$

$$\begin{aligned}
\mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) &\leq \frac{L_i^*}{b} \cdot \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) + K\sqrt{L_i^*} \log\left(L_i^*\right)^{3/4} \\
&\leq \frac{L_i^*}{b} \cdot \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) + K\sqrt{b} \log\left(b\right)^{3/4} \\
&\leq \frac{L_i^*}{b} \cdot \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/4}
\end{aligned}$$

with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/2}\right)$. Let $P_i \cap \mathcal{I}_{\geq\gamma}$ denote the instance we obtain from P_i by removing all small items. Then, it holds

$$\begin{aligned}
A'(P_i \cap \mathcal{I}_{\geq\gamma}) &\leq \frac{\min\{L_i, L_i^*\}}{L_i^*} \cdot \mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) + K\sqrt{\min\{L_i, L_i^*\}} \log\left(\min\{L_i, L_i^*\}\right)^{3/4} \\
&\leq \frac{\min\{L_i, L_i^*\}}{L_i^*} \cdot \mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) + K\sqrt{b} \log\left(b\right)^{3/4} \\
&\leq \frac{\min\{L_i, L_i^*\}}{L_i^*} \cdot \mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/4},
\end{aligned}$$

with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/2}\right)$.

Plugging both inequalities together, it follows that

$$A'(P_i \cap \mathcal{I}_{\geq\gamma}) \leq \frac{\min\{L_i, L_i^*\}}{b} \cdot \mathrm{OPT}\left(\mathcal{I}_{\geq\gamma}\right) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}.$$

Then, summing up yields

$$\begin{aligned}
\sum_{i=t}^{m} A'(P_i \cap \mathcal{I}_{\geq\gamma}) &\leq \sum_{i=t}^{m} \left(\frac{\min\{L_i, L_i^*\}}{b} \cdot \mathrm{OPT}\left(\mathcal{I}_{\geq\gamma}\right) + K\sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right) \\
&\leq \left(\frac{\mathrm{OPT}\left(\mathcal{I}_{\geq\gamma}\right)}{b} \sum_{i=t}^{m} \min\{L_i, L_i^*\}\right) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4} \\
&\leq \frac{\mathrm{OPT}\left(\mathcal{I}_{\geq\gamma}\right)}{b} \sum_{i=t}^{m} \left(\frac{|P_i|}{n} \cdot b + \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right) \\
&\qquad + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4} \\
&\leq \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) \cdot \frac{1}{n} \sum_{i=t}^{m} |P_i| + \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) \cdot K_\gamma \frac{\log(\mathrm{OPT}(\mathcal{I}))^{3/4}}{\sqrt{\mathrm{OPT}(\mathcal{I})}} \\
&\qquad + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4} \\
&\leq \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}.
\end{aligned}$$

All in all, we see that – assuming that $\mathcal{E}_L$ takes place – the number of opened bins because of large items in the phases $P_t, \ldots, P_m$ is bounded from above by $\mathrm{OPT}(\mathcal{I}_{\geq\gamma}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}$ with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log\left(\mathrm{OPT}(\mathcal{I})\right)^{3/2}\right)$.

Finally, we want to show that $\mathcal{E}_L$ takes place with sufficient probability: We apply Bernstein's inequality and use the facts that $|P_i| \leq |P_i^*| \leq n$ for all i and $\sqrt{b}\log(b)^{3/4} \leq b$. Then, we obtain for $i \in \{t, \dots, m\}$

$$\begin{aligned}\mathbb{P}\left[|L_i - \mathbb{E}[L_i]| \geq \sqrt{b}\log(b)^{3/4}\right] &\leq \exp\left(-\frac{[\sqrt{b}\log(b)^{3/4}]^2}{2 \cdot \frac{|P_i|}{n} \cdot b + \frac{2}{3}\sqrt{b}\log(b)^{3/4}}\right)\\ &\leq \exp\left(-\frac{1}{3}\log(b)^{3/2}\right)\\ &\leq \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)\end{aligned}$$

and

$$\mathbb{P}\left[|L_i^* - \mathbb{E}[L_i^*]| \geq \sqrt{b}\log(b)^{3/4}\right] \leq \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right).$$

Since the number of phases under consideration is bounded by a constant depending on γ, the statement follows by a union bound. ◀

3.2.2.6 Proof of Lemma 43

Here we will prove

▶ **Lemma 43.** *There exist universal positive constants C and α_γ such that the event*

$$\mathcal{G}^s_{start} := \left\{\sum_{i=0}^{t-1} A^s(P_i) \leq 9\gamma\,\mathrm{OPT}(\mathcal{I}) + 1\right\}$$

takes place with probability at least $1 - C\exp(-\alpha_\gamma\,\mathrm{OPT}(\mathcal{I}))$.

Furthermore, if $|P_m| < \gamma^2 n$, then also exist universal positive constants α_γ and C such that the event

$$\mathcal{G}^s_{end} := \{A^s(P_m) \leq 6\gamma\,\mathrm{OPT}(\mathcal{I}) + 1\}$$

takes place with probability at least $1 - C\exp(-\alpha_\gamma\,\mathrm{OPT}(\mathcal{I}))$.

Proof of Lemma 43. Since the small items are packed using First-Fit, we can estimate the number of opened bins for small items by the sum of small items in the first phases times a factor depending on γ. In order to do this, we show that the total size of *all* – not only the small items – drawn in P_t^* is not too large. Let X_j be the size of the j-th drawn item. Then, we have $\mathbb{E}\left[X_j^2\right] \leq \mathbb{E}[X_j] = S(\mathcal{I})/n$. Applying Bernstein's inequality, we see that

$$\begin{aligned}\mathbb{P}\left[S(P_t^*) \leq 3\gamma^2 S(\mathcal{I})\right] &\geq \mathbb{P}\left[S(P_t^*) \leq \mathbb{E}[S(P_t^*)] + \gamma^2 S(\mathcal{I})\right]\\ &\geq 1 - \exp\left(-\frac{[\gamma^2 S(\mathcal{I})]^2}{2 \cdot 2^{\lceil \log_2(\gamma^2 n)\rceil} \cdot \mathbb{E}[X_1^2] + \frac{2}{3} \cdot \gamma^2 S(\mathcal{I})}\right)\\ &\geq 1 - \exp\left(-\frac{\gamma^4 S(\mathcal{I})^2}{4\gamma^2 n \cdot \frac{S(\mathcal{I})}{n} + \gamma^2 S(\mathcal{I})}\right)\\ &\geq 1 - \exp\left(-\frac{1}{5}\gamma^2 S(\mathcal{I})\right)\\ &\geq 1 - \exp\left(-\alpha_\gamma\,\mathrm{OPT}(\mathcal{I})\right).\end{aligned}$$

Using the inequality $(1-\gamma)^{-1} \leq 1+2\gamma$ for $\gamma \leq 1/2$, we obtain that the number of opened bins due to small items up to phase P_t is upper bounded by

$$\left\lceil \frac{1}{1-\gamma} S(P_t^* \cap \mathcal{I}_{<\gamma}) \right\rceil \leq \left\lceil \frac{1}{1-\gamma} S(P_t^*) \right\rceil \leq \left\lceil \frac{3\gamma^2}{1-\gamma} S(\mathcal{I}) \right\rceil$$
$$\leq \frac{3\gamma^2}{1-\gamma} S(\mathcal{I}) + 1 \leq 9\gamma \,\mathrm{OPT}(\mathcal{I}) + 1$$

with probability at least $1 - \exp(-\alpha_\gamma \,\mathrm{OPT}(\mathcal{I}))$.

To analyze the last phase we proceed as before: Applying Bernstein's inequality we obtain

$$\begin{aligned}
\mathbb{P}\left[S(P_m) \leq 2\gamma^2 S(\mathcal{I})\right] &\geq \mathbb{P}\left[S(P_m) \leq \mathbb{E}\left[S(P_m)\right] + \gamma^2 S(\mathcal{I})\right] \\
&\geq 1 - \exp\left(- \frac{[\gamma^2 S(\mathcal{I})]^2}{2\gamma^2 n \cdot \frac{S(\mathcal{I})}{n} + \frac{2}{3}\gamma^2 S(\mathcal{I})} \right) \\
&\geq 1 - \exp\left(-\frac{1}{3}\gamma^2 S(\mathcal{I}) \right).
\end{aligned}$$

Hence, in this case the number of opened bins because of small items is upper bounded by $6\gamma \,\mathrm{OPT}(\mathcal{I}) + 1$ with proability at least $1 - C\exp(-\alpha_\gamma \,\mathrm{OPT}(\mathcal{I}))$. ◀

3.2.2.7 Proof of Lemma 44

Here we will prove

▶ **Lemma 44.** *There exist universal positive constants α_γ, C_γ and K_γ such that the event $\mathcal{G}^s_{remaining}$ defined as*

$$\left\{ \sum_{i=t}^{m} A^s(P_i) \leq [S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})]^+ + 7\gamma \,\mathrm{OPT}(\mathcal{I}) + K_\gamma \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4} \right\}$$

takes place with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$.

Now we study the number of opened bins in the remaining phases that are especially dedicated for small items. That is we investigate how often it will happen that we cannot put a small item into a bin that was opened in a previous phase or in space generated by a virtual item with rank one from the current phase.

The number of opened bins for small items in phase P_i depends on

- the empty space in bins opened in previous phases;
- the small items arriving in phase P_i and
- the virtual items generated in phase P_i.

We can assume without loss of generality that all bins opened in previous phases are full, so that it is not possible to put small items in these bins.

In the following we will consider the process of packing the small items in phase $P_i = (A_{2^i+1}, \ldots, A_{\min\{2\cdot 2^i,\ldots,n\}})$ in more details: For each of the first $L_i - \min\{L_i, L_i^*\}$ many large items arriving a virtual item is created, and their sizes stem from the set $\{v_1, \ldots, v_{L_i^*}\} = P_i^* \cap \mathcal{I}_{\geq\gamma}$. Now we introduce corresponding space-items: We remember that at the beginning of phase P_i a model-packing $\mathcal{M}(P_i^*)$ is computed. Let $f = f_{\mathcal{M}} : P_i^* \cap \mathcal{I}_{\geq\gamma} \to [\mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma})]$

be the function, which assigns to each item the index of its corresponding bin in the model-packing, and r the corresponding rank-function. Now, for each possible size of a virtual item v_j the corresponding space-item has size

$$e_j := \begin{cases} \min\left\{-\left(1-\sum_{m\,:\,f(v_m)=f(v_j)} v_m\right)+\gamma, 0\right\} & \text{if } r(v_j)=1, \\ 0 & \text{otherwise.} \end{cases}$$

This is motivated as follows: Space for small items can only be generated by virtual items of rank one. Furthermore, it is possible that a fraction arbitrary close to γ will not be used by small items since they do not fit into this space. We set $E_i := \{e_1, \ldots, e_{L_i^*}\}$.

Now we apply the following transformation to P_i:

- At first we delete the last $L_i - \min\{L_i, L_i^*\}$ large items from P_i.
- Then, we replace all remaining large items by the corresponding virtual items, and afterwards the virtual items by the corresponding empty-space items.

Let $P_i^r = (R_1, \ldots, R_{|P_i|-(L_i-\min\{L_i,L_i^*\})})$ be the result of this transformation and moreover we set $R_0 := 0$.

If we look at the partial sums of P_i^r we observe that every time a new small item arrives we add some “demand” and every time a virtual item of rank one arrives we generate some empty space. So this process is closely related to the number of bins we need to open for small items as the following lemma will show:

▶ **Lemma 45.** *Set $M_i := \max_{\{0\le u\le |P_i|-(L_i-\min\{L_i,L_i^*\})\}} \sum_{m=1}^{u} R_m$. Then, the number of opened bins for small items in phase P_i is bounded from above by $(1+2\gamma)M_i + 1$.*

Proof. We have

$$\left\lceil \frac{1}{1-\gamma} \max_u \sum_{m=1}^{u} R_m \right\rceil \le (1+2\gamma) \max_u \sum_{m=1}^{u} R_m + 1.$$

Here, the factor $(1-\gamma)^{-1}$ stems from the fact that in the worst case for each bin that is designated fully for small items space of size nearly γ is wasted. ◀

Now we want to take a closer look at P_i^r: Let $\mathcal{Q}$ denote the multiset of all values, which arise in P_i^r. Then, we observe the following two things:

- $\mathcal{Q}$ is a subset of $\mathcal{I}_{<\gamma} \cup E_i$, and
- each realization of P_i^r with all values in $\mathcal{Q}$ is equally likely.

Let $\hat{\mathcal{Q}}$ be the set containing all possible candidate sets $\mathcal{Q}$, which can arise with positive probability. Moreover, let p denote the probability distribution on $\hat{\mathcal{Q}}$, where $p(\mathcal{Q})$ is the probability that P_i^r is equal to a permutation of the values in $\mathcal{Q}$. Then, we can look at P_i^r in the following way: At first we draw a random set $\mathcal{Q}$ from $\hat{\mathcal{Q}}$ according to p. Then, we generate a random permutation of the items in $\mathcal{Q}$ and call this $I_{\mathcal{Q}}^{\sigma}$.

The remaining steps of the proof are the following:

- At first we show that for a set $\mathcal{Q}$ the maximum of the partial sums of a random permutation of the values in $\mathcal{Q}$ is not much larger than the total size of all values in $\mathcal{Q}$.
- Then, we will show that the total size of all items in $\mathcal{Q}$, drawn according to p, is not too large.

▶ **Lemma 46.** *Let $\mathcal{Q} \in \hat{\mathcal{Q}}$, and let $I^\sigma_{\mathcal{Q}}$ be a random permutation of the elements in $\mathcal{Q}$. Moreover, let $M(I^\sigma_{\mathcal{Q}})$ denote the maximum of the partial sums. Then, there exist constants C and α_γ such that*

$$M(I^\sigma_{\mathcal{Q}}) \leq [S(\mathcal{Q})]^+ + \sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}$$

with probability at least $1 - C\exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$.

▶ **Lemma 47.** *Let $\mathcal{Q}$ be randomly drawn from $\hat{\mathcal{Q}}$ according to p. Then, there exist universal constants α_γ, C and K_γ such that*

$$\begin{aligned}&S(\mathcal{Q})\\ &\leq \frac{|P_i|}{n}\cdot[S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})]^+ + \frac{|P_i|}{n}\cdot\gamma\,\mathrm{OPT}(\mathcal{I}) + K_\gamma\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\end{aligned} \tag{8}$$

with probability at least $1 - C\exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$.

Combining the statements given in Lemma 45, 46 and 47 we see that with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$ the number of opened bins for small items is bounded from above by

$$\begin{aligned}&(1+2\gamma)\max_{0\leq u\leq|\mathcal{Q}|}\sum_{j=1}^{u} X_j + 1\\ &\leq (1+2\gamma)\left[\frac{|P_i|}{n}\left(S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})\right)^+ + \frac{|P_i|}{n}\gamma\,\mathrm{OPT}(\mathcal{I})\right.\\ &\qquad\left. + K_\gamma\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right] + 1\\ &\leq \frac{|P_i|}{n}[S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})]^+ + \frac{|P_i|}{n}7\gamma\,\mathrm{OPT}(\mathcal{I}) + K_\gamma\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}.\end{aligned}$$

All in all, applying a union bound we see that the number of opened bins for small items in the phases $t, \ldots, m$ is bounded from above by

$$\begin{aligned}&\sum_{j=t}^{m}\left[\frac{|P_i|}{n}[S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})]^+ + \frac{|P_i|}{n}7\gamma\,\mathrm{OPT}(\mathcal{I}) + K_\gamma\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right]\\ &\leq [S(\mathcal{I}) - \mathrm{OPT}(\mathcal{I}_{\geq\gamma})]^+ + 7\gamma\,\mathrm{OPT}(\mathcal{I}) + K_\gamma\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\end{aligned}$$

with probability at least $1 - C_\gamma \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$. It remains to prove Lemma 46 and Lemma 47.

We start with the proof of Lemma 46.

Proof of Lemma 46. The proof is based on an application of the maximal inequality for sampling without replacement. We can assume without loss of generality that $S(\mathcal{Q}) \geq 0$, otherwise we increase the size of space-items (that is we decrease the space they offer). Let $I^\sigma_{\mathcal{Q}} = (X_1, \ldots, X_{|\mathcal{Q}|})$ with $X_i = Q_{\sigma(i)}$ be a random permutation of the elements in $\mathcal{Q}$. We set $\tilde{X}_j := X_j - \frac{S(\mathcal{Q})}{|\mathcal{Q}|}$. Since the random variables are only shifted by a constant they are still exchangeable. Applying Proposition 14 we obtain

$$\mathbb{P}\left[\max_{1\leq u\leq|\mathcal{Q}|}\left|\sum_{j=1}^{u}\tilde{X}_j\right| > \sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right]$$

$$\leq C \cdot \mathbb{P}\left[\left|\sum_{j=1}^{\lfloor|\mathcal{Q}|/2\rfloor} \tilde{X}_j\right| > \frac{1}{C}\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right].$$

Assume that

$$\max_{1\leq u\leq|\mathcal{Q}|}\left|\sum_{j=1}^{u}\tilde{X}_j\right| \leq \sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}.$$

Then, we have for all $u \in \{1,\dots,|\mathcal{Q}|\}$

$$\sum_{j=1}^{u} X_j \leq \frac{u}{|\mathcal{Q}|}\cdot[S(\mathcal{Q})]^+ + \sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}$$
$$\leq [S(\mathcal{Q})]^+ + \sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}.$$

So, to show the statement of the lemma it suffices to bound

$$\mathbb{P}\left[\left|\sum_{j=1}^{\lfloor|\mathcal{Q}|/2\rfloor} \tilde{X}_j\right| > \frac{1}{C}\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right]$$

from above. In order to do this, we will apply Bernstein's inequality, that is, we will bound the term

$$\exp\left(-\frac{\mathrm{OPT}(\mathcal{I})\log(\mathrm{OPT}(\mathcal{I}))^{3/2}}{2C^2\sum_{j=1}^{\lfloor|\mathcal{Q}|/2\rfloor}\mathbb{E}\left[\tilde{X}_j^2\right]+\frac{2}{3}M\cdot C\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}}\right)$$
$$\leq \exp\left(-\frac{\mathrm{OPT}(\mathcal{I})\log(\mathrm{OPT}(\mathcal{I}))^{3/2}}{C^2\cdot|\mathcal{Q}|\cdot\mathbb{E}\left[\tilde{X}_1^2\right]+\frac{2}{3}M\cdot C\sqrt{\mathrm{OPT}(\mathcal{I})}\log(\mathrm{OPT}(\mathcal{I}))^{3/4}}\right)$$

from above.

At first we want to give an upper bound M for $|\tilde{X}_i|$. We obtain

$$|\tilde{X}_i| = \left|X_i - \frac{S(\mathcal{Q})}{|\mathcal{Q}|}\right| \leq |X_i| + \frac{|S(\mathcal{Q})|}{|\mathcal{Q}|}$$
$$\leq 1+\max\left\{\frac{|S(\mathcal{Q}\cap\mathcal{I}_{<\gamma})|}{|\mathcal{Q}\cap\mathcal{I}_{<\gamma}|},\frac{|S(\mathcal{Q}\cap E_i)|}{|\mathcal{Q}\cap E_i|}\right\} \leq 1+\max\{\gamma,1\}\leq 2.$$

Since the absolute value of every element in $\mathcal{Q}$ is bounded from above by 1 it follows that $|S(\mathcal{Q})|/|\mathcal{Q}| \leq 1$. Hence, it follows that

$$\mathbb{E}\left[\tilde{X}_1^2\right] = \mathbb{E}\left[\left(X_1 - \frac{S(\mathcal{Q})}{|\mathcal{Q}|}\right)^2\right]$$
$$\leq \mathbb{E}\left[X_1^2\right] + 2\mathbb{E}\left[|X_1|\cdot\frac{|S(\mathcal{Q})|}{|\mathcal{Q}|}\right] + \frac{1}{|\mathcal{Q}|}\mathbb{E}\left[\frac{|S(\mathcal{Q})|}{|\mathcal{Q}|}\cdot|S(\mathcal{Q})|\right]$$
$$\leq \mathbb{E}\left[|X_1|\right] + 2\mathbb{E}\left[|X_1|\right] + \frac{1}{|\mathcal{Q}|}\cdot\mathbb{E}\left[|S(\mathcal{Q})|\right]$$
$$\leq 3\mathbb{E}\left[|X_1|\right] + \frac{1}{|\mathcal{Q}|}\cdot|\mathcal{Q}|\cdot\mathbb{E}\left[|X_1|\right]$$
$$\leq 4\mathbb{E}\left[|X_1|\right].$$

Finally, we want to show that we can bound $|\mathcal{Q}| \cdot \mathbb{E}[|X_1|]$ in terms of $\mathrm{OPT}(\mathcal{I})$. We obtain

$$\begin{aligned}|\mathcal{Q}| \cdot \mathbb{E}[|X_1|] &= |\mathcal{Q}| \cdot \left(\frac{|\mathcal{Q} \cap \mathcal{I}_{<\gamma}|}{|\mathcal{Q}|} \cdot \frac{S(\mathcal{I}_{<\gamma})}{|\mathcal{I}_{<\gamma}|} + \frac{|\mathcal{Q} \cap E_i|}{|\mathcal{Q}|} \cdot 1 \right) \\ &\leq S(\mathcal{I}_{<\gamma}) + |\mathcal{I}_{\geq\gamma}| \\ &\leq \left(1 + \frac{1}{\gamma}\right) \mathrm{OPT}(\mathcal{I}).\end{aligned}$$

Combining the bounds shows that there exist constants C and α_γ such that the maximum of the partial sums of a random permutation of the elements in $\mathcal{Q}$ is bounded from above by $[S(\mathcal{Q})]^+ + \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}$ with probability at least $1 - \exp\left(-\alpha_\gamma \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right)$. ◀

Proof of Lemma 47. The proof is a bit longish and based on several applications of concentration bounds. It holds $S(\mathcal{Q}) = S(\mathcal{Q} \cap \mathcal{I}_{<\gamma}) + S(\mathcal{Q} \cap E_i)$. We will investigate both terms independently:

- $S(\mathcal{Q} \cap \mathcal{I}_{<\gamma})$ can be controlled using Bernstein's inequality.
- Estimating $S(\mathcal{Q} \cap E_i)$ is more tedious as we have to estimate at first $S(E_i)$ and afterwards the total size of the items in $\mathcal{Q} \cap E_i$.

▶ **Lemma 48.** *We have*

$$\mathbb{P}\left[S(P_i \cap \mathcal{I}_{<\gamma}) \geq \frac{|P_i|}{n} \cdot S(\mathcal{I}_{<\gamma}) + \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right] \leq \exp\left(-\frac{1}{3} \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right).$$

Proof. This bound is just an application of Bernstein's inequality: Let Y_j be equal to zero if the j-th drawn item is large and otherwise the size of the drawn small item. Then, we have $\sum_{j=1}^{|P_i|} Y_j = S(P_i \cap \mathcal{I}_{<\gamma})$ and $\mathbb{E}[S(P_i \cap \mathcal{I}_{<\gamma})] = \frac{|P_i|}{n} \cdot S(\mathcal{I}_{<\gamma})$. It follows that

$$\begin{aligned}&\mathbb{P}\left[S(P_i \cap \mathcal{I}_{<\gamma}) \geq \mathbb{E}[S(P_i \cap \mathcal{I}_{<\gamma}] + \sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I}))^{3/4}\right] \\ &\leq \exp\left(-\frac{\mathrm{OPT}(\mathcal{I}) \log(\mathrm{OPT}(\mathcal{I}))^{3/2}}{2|P_i| \cdot \mathbb{E}[X_1^2] + \frac{2}{3}\sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I})^{3/4}}\right) \\ &\leq \exp\left(-\frac{\mathrm{OPT}(\mathcal{I}) \log(\mathrm{OPT}(\mathcal{I}))^{3/2}}{2|P_i| \cdot \mathbb{E}[X_1] + \frac{2}{3}\sqrt{\mathrm{OPT}(\mathcal{I})} \log(\mathrm{OPT}(\mathcal{I})^{3/4}}\right) \\ &\leq \exp\left(-\frac{\mathrm{OPT}(\mathcal{I}) \log(\mathrm{OPT}(\mathcal{I}))^{3/2}}{2|P_i| \cdot \frac{S(\mathcal{I})}{n} + \mathrm{OPT}(\mathcal{I})}\right) \\ &\leq \exp\left(-\frac{1}{3} \log(\mathrm{OPT}(\mathcal{I}))^{3/2}\right).\end{aligned}$$

◀

Now, we want to control $S(\mathcal{Q} \cap E_i)$. We remember that

$$S(E_i) \leq S(P_i^* \cap \mathcal{I}_{\geq\gamma}) - (1-\gamma) \mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}). \tag{9}$$

In the following two lemmata we will give bounds for $S(P_i^* \cap \mathcal{I}_{\geq\gamma})$ and $\mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma})$:

▶ **Lemma 49.** *It holds*

$$\mathbb{P}\left[S(P_i^* \cap \mathcal{I}_{\geq\gamma}) \leq \frac{L_i^*}{|\mathcal{I}_{\geq\gamma}|} \cdot S(\mathcal{I}_{\geq\gamma}) + \sqrt{L_i^*}\log(L_i^*)^{3/4} \mid L_i^*\right] \geq 1 - \exp\left(-\frac{1}{3}\log(L_i^*)^{3/2}\right).$$

Proof. The proof of this statement is just another application of Bernstein's inequality. Then, we obtain

$$\begin{aligned}
&\mathbb{P}\left[S(P_i^* \cap \mathcal{I}_{\geq\gamma}) \leq \frac{L_i^*}{|\mathcal{I}_{\geq\gamma}|} \cdot S(\mathcal{I}_{\geq\gamma}) + \sqrt{L_i^*}\log(L_i^*)^{3/4} \mid L_i^*\right] \\
\geq& 1 - \exp\left(-\frac{L_i^*\log(L_i^*)^{3/2}}{2L_i^* \cdot \frac{S(\mathcal{I}_{\geq\gamma})}{|\mathcal{I}_{\geq\gamma}|} + \frac{2}{3}\sqrt{L_i^*}\log(L_i^*)^{3/4}}\right) \\
\geq& 1 - \exp\left(-\frac{L_i^*\log(L_i^*)^{3/2}}{2L_i^* + \frac{2}{3}\sqrt{L_i^*}\log(L_i^*)^{3/4}}\right) \\
\geq& 1 - \exp\left(-\frac{\log(L_i^*)^{3/2}}{2 + \frac{2}{3}\frac{\log(L_i^*)^{3/4}}{\sqrt{L_i^*}}}\right) \\
\geq& 1 - \exp\left(-\frac{1}{3}\log(L_i^*)^{3/2}\right).
\end{aligned}$$

◀

▶ **Lemma 50.** *Assume that there exists a constant β_γ depending on γ such that $L_i^* \geq \beta_\gamma|\mathcal{I}_{\geq\gamma}|$. Then, there exist universal constants C_γ, K and α such that*

$$\mathbb{P}\left[\mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) \geq \frac{L_i^*}{|\mathcal{I}_{\geq\gamma}|} \cdot \mathrm{OPT}(\mathcal{I}_{\geq\gamma}) - K\sqrt{L_i^*}\log(L_i^*)^{3/4} \mid L_i^*\right] \geq 1 - C_\gamma \exp\left(-\alpha\log(L_i^*)^{3/2}\right).$$

Proof. We divide the input into several phases of length L_i^* (apart from the last phase, which could be smaller). Then, we use the packing of the first phase (which is identically distributed to the distribution of the packing of the items in $P_i^* \cap \mathcal{I}_{\geq\gamma}$) as a model-packing for the items of the following phases. Then, it follows as in the analysis of the auxiliary algorithm for cardinality-constrained bin packing that with high probability we can pack the items from phase two in at most

$$\mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) + K\sqrt{L_i^*}\log(L_i^*)^{3/4}$$

many bins. Applying the same argument for the remaining phases, using a union bound and the fact that $L_i^* \geq \beta_\gamma|\mathcal{I}_{\geq\gamma}|$ it follows that

$$\mathrm{OPT}(P_i^* \cap \mathcal{I}_{\geq\gamma}) \geq \frac{L_i^*}{|\mathcal{I}_{\geq\gamma}|}\mathrm{OPT}(\mathcal{I}_{\geq\gamma}) - K\sqrt{L_i^*}\log(L_i^*)^{3/4}$$

with probability at least $1 - C_\gamma \exp\left(-\alpha\log(L_i^*)^{3/2}\right)$. ◀

After giving bounds for both terms on the right hand side in (9), we are finally able to control the generated space by the actually created virtual items:

▶ **Lemma 51.** *It holds*

$$\mathbb{P}\left[S(\mathcal{Q} \cap E_i) \leq \min\{L_i, L_i^*\} \cdot \frac{S(E_i)}{L_i^*} + \sqrt{L_i}\log(L_i)^{3/4} \mid L_i^*, L_i\right] \geq 1 - \exp\left(-\frac{1}{3}\log(L_i)^{3/2}\right).$$

Proof. This proof is again a direct application of Bernstein's inequality and exploiting the fact that $S(E_i) \leq L_i^*$. ◀

Eventually, Lemma 47 follows by combining the statements of Lemma 48, 49, 50 and 51 with concentration inequalities for L_i and L_i^* and using a union bound. ◀

3.2.3 A Lower Bound for Classical Bin Packing in the Partial-permutations Model

Here, we will prove:

▶ **Proposition 35.** *For online classical bin packing there exists for every $p \in [0,1)$ an ϵ_p greater than zero such that for all deterministic online algorithms A it holds $\mathrm{RR}_p(A) \geq 1+\epsilon_p$.*

3.2.3.1 High-Level Proof

The first important thing to notice is that the performance measure $\mathrm{RR}_p(A)$ is more pessimistic than $\mathrm{RR}(A)$:

▶ **Lemma 52.** *Let $p \in [0,1)$. Then, for an arbitrary deterministic algorithm A it is true that $\mathrm{RR}(A) \leq \mathrm{RR}_p(A)$.*

Now, we define for $n \in \mathbb{N}$ the instance $I = (3/7, \dots, 3/7, 4/7, \dots, 4/7)$ consisting of $2n$ items of size $3/7$ in the beginning, followed by $2n$ items of size $4/7$. It is clear that an optimal solution for the first $2n$ items needs n many bins, and that $\mathrm{OPT}(I) = 2n$. Moreover, it is straightforward to show that there is no way to pack the items online such that we obtain an optimal solution after packing the first $2n$ items and in the end simultaneously.

Let $p \in (0,1)$ and A be a deterministic online algorithm. We denote by $I^{p,\sigma}$ a random instance that we obtain from I by performing the random perturbation described in the partial-permutations model.

The idea of the proof is the following: We will show that for each n large enough there exists an ϵ_p greater than zero such that either $\frac{\mathbb{E}[A(I^{p,\sigma})]}{\mathrm{OPT}(I)} \geq 1+\epsilon_p$ or that we can find an instance $\tilde{I}$ consisting of $2n$ items with $\mathrm{OPT}(\tilde{I}) = n$ such that $\frac{\mathbb{E}[A(\tilde{I}^\sigma)]}{\mathrm{OPT}(\tilde{I})} \geq 1+\epsilon_p$. Here, $\tilde{I}^\sigma$ denotes a random permutation of the items in $\tilde{I}$ as in the random-order model. Then, Proposition 35 follows from Lemma 52.

We assume that

$$\mathbb{E}\left[A(I^{p,\sigma})\right] \leq \left(1 + \frac{1}{12}\cdot(1-p)\right)2n = \left(1+\frac{1}{12}\cdot(1-p)\right)\mathrm{OPT}(I). \tag{10}$$

If this is not the case, we have already found an instance I with $\frac{\mathbb{E}[A(I^{p,\sigma})]}{\mathrm{OPT}(I)} > 1 + \frac{1}{12}(1-p)$. So let $\mathcal{F} = \left\{I^{p,\sigma} : A(I^{p,\sigma}) \leq \left(1+\frac{1}{6}(1-p)\right)\cdot 2n\right\}$. If condition (10) is satisfied, we have $\mathbb{P}\left[\mathcal{F}\right] \geq 1/2$, otherwise it would hold that

$$\mathbb{E}\left[A(I^{p,\sigma})\right] > \mathbb{P}\left[\mathcal{F}\right]\cdot 2n + (1-\mathbb{P}\left[\mathcal{F}\right])\cdot\left(1+\frac{1}{6}(1-p)\right)\cdot 2n$$

$$> \frac{1}{2} \cdot 2n + \frac{1}{2} \cdot \left(1 + \frac{1}{6}(1-p)\right) \cdot 2n$$
$$= \left(1 + \frac{1}{12}(1-p)\right) \cdot 2n.$$

Moreover, for $I^{p,\sigma}$ let $M^{p,\sigma}$ denote the number of items of size $4/7$ among the first $2n$ items. We set $\mathcal{E} = \left\{I^{p,\sigma} : M^{p,\sigma} \leq \frac{1}{2}(1+p)n\right\}$. The following lemma shows that it is not possible that A performs well on realizations in $\mathcal{E}$ at the time points $2n$ and $4n$ simultaneously:

▶ **Lemma 53.** *Let $p' \in [0,1)$. Let I be an instance consisting of $2n$ items of size $3/7$ and $2n$ items of size $4/7$, where the number of items of size $4/7$ among the first $2n$ items is bounded from above by $p' \cdot n$. Then, no online algorithm A can be better than $(1+(1-p')/3)$-competitive after packing the first $2n$ items from I and packing all items of I.*

Furthermore, the event $\mathcal{E}$ takes place with high probability as the following statement shows:

▶ **Lemma 54.** *Let I be an instance as described consisting of $4n$ many items. Assume that $I^{p,\sigma}$ is a partially permuted instance and let $M^{p,\sigma}$ denote the number of items of size $4/7$ among the first $2n$ many items in $I^{p,\sigma}$. Then, there exist positive constants α_p and C_p such that*

$$\mathbb{P}\left[M^{p,\sigma} \geq (1+p)/2 \cdot n\right] \leq C_p \exp\left(-\alpha_p n\right).$$

Using a union bound, it follows from Lemma 54 that there exist positive constants C_p and α_p such that we have for n large enough

$$\mathbb{P}\left[\mathcal{E} \cap \mathcal{F}\right] \geq 1/2 - C_p \exp\left(-\alpha_p n\right). \tag{11}$$

Now, for an instance $I^{p,\sigma} = (A_1, \ldots, A_{4n})$ let g denote the projection onto the first $2n$ coordinates, that is we have $g(I^{\sigma,p}) = (A_1, \ldots, A_{2n})$. Set $p' = 1/2 \cdot (1+p)$ and let $I^{p,\sigma}$ be a realization in $\mathcal{E} \cap \mathcal{F}$. Then, $I^{p,\sigma}$ satisfies the following properties: At first we notice that $M^{p,\sigma} \leq p'n \leq n$, and therefore also $\mathrm{OPT}(g(I^{p,\sigma})) = n$. Then, we observe that

$$A(I^{p,\sigma}) \leq \left(1 + \frac{1}{6}\left(1-p\right)\right) 2n = \left(1 + \frac{1}{3}\left(1-p'\right)\right) 2n.$$

Hence, it follows from Lemma 53 that

$$A(g(I^{p,\sigma})) > \left(1 + \frac{1}{3}\left(1-p'\right)\right) \cdot n.$$

So, using (11) we obtain the following estimate:

$$\mathbb{E}\left[A(g(I^{p,\sigma}))\right] \geq \mathbb{P}\left[\mathcal{E} \cap \mathcal{F}\right] \cdot \left[1 + \frac{1}{3}(1-p')\right] n + (1 - \mathbb{P}\left[\mathcal{E} \cap \mathcal{F}\right])n$$
$$\geq \left(\frac{1}{2} - C_p \exp\left(-\alpha_p n\right)\right) \cdot \left[\frac{7}{6} - \frac{1}{6}p\right] n + \left(\frac{1}{2} + C_p \exp\left(-\alpha_p n\right)\right) \cdot n$$
$$\geq \left[\frac{13}{12} - \frac{1}{12}p\right] n - C_p \exp\left(-\alpha_p n\right) n.$$

Now let $\mathcal{H}$ denote all possible projections onto the first $2n$ coordinates, that is we have $\mathcal{H} = \{g(I^{p,\sigma})\}$. Then we observe that projections that contain the same multiplicities of item types are equally likely:

▶ **Observation 55.** *Let $I', I'' \in \mathcal{H}$ such that the number of items of size $4/7$ in I' and I'' is equal. Then $\mathbb{P}\left[g(I^{p,\sigma}) = I'\right] = \mathbb{P}\left[g(I^{p,\sigma}) = I''\right]$.*

Moreover, for $i \in \{0, \ldots, 2n\}$ let I_i denote the instance beginning with i many items of size $4/7$, followed by $2n - i$ many items of size $3/7$. Then, it follows by the previous observation that there exists a probability mass function $q : \{0, \ldots, 2n\} \to [0, 1]$ such that

$$\mathbb{E}\left[A(g(I^{p,\sigma}))\right] = \sum_{i=0}^{2n} q_i \cdot \mathbb{E}\left[A(I_i^{\sigma})\right].$$

Furthermore, it follows again from Lemma 54 that we have

$$\sum_{i=0}^{\left\lfloor \frac{1}{2}(1-p)n \right\rfloor} q_i \cdot \mathbb{E}\left[A(I_i^{\sigma})\right] \geq \mathbb{E}\left[A(g(I^{p,\sigma}))\right] - C_p \exp\left(-\alpha_p \cdot n\right) \cdot 4n.$$

Hence, we obtain

$$\begin{aligned} \sum_{i=0}^{\left\lfloor \frac{1}{2}(1-p)n \right\rfloor} q_i \cdot \mathbb{E}\left[A(I_i^{\sigma})\right] &\geq \left[\frac{13}{12} - \frac{1}{12}p\right] n - 5C_p \exp\left(-\alpha_p n\right) n \\ &= \left[1 + \frac{1}{12}(1-p)\right] n - 5C_p \exp\left(-\alpha_p n\right) n. \end{aligned}$$

Therefore, for n large enough there must exist an $i \in \left\{0, \ldots, \left\lfloor \frac{1}{2}(1-p)n \right\rfloor\right\}$ such that $\frac{\mathbb{E}[A(I_i^{\sigma})]}{\mathrm{OPT}(I_i)} \geq \left[1 + \frac{1}{24}(1-p)\right]$. This shows the statement.

3.2.3.2 Proof of Lemma 52

Here we will prove

▶ **Lemma 52.** *Let $p \in [0, 1)$. Then, for an arbitrary deterministic algorithm A it is true that $\mathrm{RR}(A) \leq \mathrm{RR}_p(A)$.*

Proof of Lemma 52. For a set of items $\mathcal{I}$ let $I_1, \ldots, I_{|\mathcal{I}|!}$ denote all possible permutations of the items. The proof then is a consequence of the following observation: If we choose a random permutation I_i with respect to the uniform distribution, and afterwards perturb this permutation according to the partial-permutations model, then we obtain again a permutation that is distributed according to the uniform distribution. That is, we have

$$\mathbb{E}\left[A(I^{\sigma})\right] = \frac{1}{|\mathcal{I}|!} \sum_{i=1}^{|\mathcal{I}|!} \mathbb{E}\left[A(I_i^{p,\sigma})\right].$$

Then, it follows that there exists at least one permutation I_i with $\mathbb{E}\left[A(I_i^{\sigma,p})\right] \geq \mathbb{E}\left[A(I^{\sigma})\right]$. Hence, we obtain $\mathrm{RR}_p(A) \geq \mathrm{RR}(A)$. ◀

3.2.3.3 Proof of Lemma 53

Here we will prove

▶ **Lemma 53.** *Let $p' \in [0, 1)$. Let I be an instance consisting of $2n$ items of size $3/7$ and $2n$ items of size $4/7$, where the number of items of size $4/7$ among the first $2n$ items is bounded from above by $p' \cdot n$. Then, no online algorithm A can be better than $(1+(1-p')/3)$-competitive after packing the first $2n$ items from I and packing all items of I.*

Proof of Lemma 53. Let $I = (a_1, \ldots, a_{2n}, a_{2n+1}, \ldots, a_{4n})$ be an instance for the problem satisfying the stated conditions. Let M denote the number of items of size $4/7$ among the first $2n$ items in I. Then, every algorithm A has to pack those items into separate bins. Moreover, the remaining $2n - M$ items of size $3/7$ must be packed. It follows from our assumption that $2n - M \geq M$. We can assume without loss of generality that A packs M items of size $3/7$ together with the items of size $4/7$. Hence, we have M perfectly packed bins. For the first half, it remains to pack the remaining $2n - 2M$ items of size $3/7$. Let Q denote the number of additionally opened bins containing two items of size $3/7$. Then, $2n - 2M - 2Q$ is the number of bins containing a single item of size $3/7$. All in all, after packing the first $2n$ items we have opened M bins containing an item of size $4/7$ and $3/7$, Q bins containing two items of size $3/7$ and $2n - 2M - 2Q$ bins containing only one item of size $3/7$. So, the total number of opened bins by the algorithm is equal to $2n - M - Q$. Moreover, an optimal solution can pack the first $2n$ items into n bins.

Now, we consider the packing of the items of the second half. We know that there are M items of size $3/7$ and $2n - M$ items of size $4/7$. Without loss of generality $2n - 2M - 2Q$ many of the items of size $4/7$ can be packed in the open bins from the first half. Hence, in the second half we have to open another $M + 2Q$ many bins for the items of size $4/7$ and can pack the remaining items of size $3/7$ into those bins. So, the total number of opened bins after packing the whole instance by A is equal to $2n + Q$. Furthermore, an optimal packing of I needs $2n$ many bins.

A short calculation shows that the expression $\max\{(2n - M - Q)/n, (2n + Q)/(2n)\}$ is minimized for $Q = 2(n - M)/3$. Then, it follows that no algorithm can be better than

$$\frac{2n + Q}{2n} = \frac{2n + \frac{2}{3}(n - M)}{2n} = 1 + \frac{1}{3}\left(1 - \frac{M}{n}\right) \geq 1 + \frac{1}{3}\left(1 - \frac{p'n}{n}\right) = 1 + \frac{1}{3}(1 - p')$$

competitive. ◀

3.2.3.4 Proof of Lemma 54

Here we will prove

▶ **Lemma 54.** *Let I be an instance as described consisting of $4n$ many items. Assume that $I^{p,\sigma}$ is a partially permuted instance and let $M^{p,\sigma}$ denote the number of items of size $4/7$ among the first $2n$ many items in $I^{p,\sigma}$. Then, there exist positive constants α_p and C_p such that*

$$\mathbb{P}\left[M^{p,\sigma} \geq (1 + p)/2 \cdot n\right] \leq C_p \exp\left(-\alpha_p n\right).$$

Proof of Lemma 54. Let H_1 denote the set of selected indices for the partial permutation in the first half, and H_2 the selected indices in the second half. Then, we have $\mathbb{E}\left[|H_1|\right] = \mathbb{E}\left[|H_2|\right] = 2n \cdot p$. Since the indices are chosen independently and identically distributed with probability p, we can apply concentration bounds for $|H_1|$ and $|H_2|$.

Furthermore, if we fix sets H_1 and H_2, then the expected number of items of size $4/7$ among the first $2n$ many items after choosing a random permutation is $\frac{|H_1| \cdot |H_2|}{|H_1| + |H_2|}$. Moreover, using Hoeffding's inequality we obtain that $M^{p,\sigma}$ for fixed H_1 and H_2 is also highly concentrated around its mean. Combining both concentration bounds and using a union bound we obtain the desired result. ◀

3.2.4 Existence of an Optimal Algorithm for Class-constrained Bin Packing with respect to the Average Performance Ratio

Here, we will prove:

▶ **Proposition 36.** *Consider online class-constrained bin packing with parameter k, and let $\mathcal{D}$ denote the set of all distributions on $[0,1] \times \mathbb{N}$. Then, there exists a randomized algorithm A with* $\mathrm{APR}(A, \mathcal{D}) = 1$.

3.2.4.1 Description of the Algorithm and High-Level Proof

Let F be a distribution on $[0,1] \times \mathbb{N}$. We can think of F as a collection of distributions F_{color}, F_1, F_2, ... Here, F_{color} is a distribution on $\mathbb{N}$ and F_i (for $i \in \mathbb{N}$) are distributions on $[0,1]$. We can think of generating a random item (S, C) according to F as follows: At first we generate a random color C distributed according to F_{color}, and afterwards we generate the size S distributed according to F_C.

We see that we explicitly allow items of size zero. Packing these items in classical bin packing is trivial, but in class-constrained bin packing they can matter. Our algorithm A treats items of size equal to zero and greater than zero differently. For items of size zero we proceed as follows: If already an item of the same color arrived, then we place the new item in the same bin. Otherwise, we put the item into the first possible bin that contains only items of size zero, and if also such a bin does not exist we open a new bin for the item. We notice that this procedure is optimal in the case that all items have size zero.

Let $A^0(I_n^F)$ denote the number of opened bins for items of size zero. We will show that this term is negligible if we draw items with size greater than zero with positive probability.

▶ **Lemma 56.** *Let F be a distribution, and let S be the size of a random item drawn according to F. If $\mathbb{P}[S > 0] > 0$ then it holds true that*

$$\limsup_{n\to\infty} \mathbb{E}\left[\frac{A^0(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] = 0.$$

Hence, we will make the following assumption:

▶ **Assumption 3.** *Let F be a distribution and let S be the size of a random item drawn according to F. Then, it holds $\mathbb{P}[S > 0] = 1$.*

For items of size greater than zero we apply the approach from [72]:

- We generate virtual items in the same way as seen in the part about cardinality-constrained and classical bin packing and pack them using an auxiliary algorithm. There are two changes in the auxiliary algorithm: We only replace second-order virtual items of the same color and the same rank. Moreover, we pack items that cannot replace a placeholder using Next-Fit for each color separately.
- Moreover, if the arriving real item is of color c we try to replace the smallest possible virtual item of color c by the real item.

We see that in the replacement procedure (in the auxiliary algorithm and when we replace virtual items by real items) we perform this action for each color separately. It is unlikely that using this approach we obtain an error bound comparable to (5), because for distributions F with a heavy-tailed distribution on the colors F_{color} there will be a significant amount of items with colors that do not appear in the model-packing, but the error will be still sublinear. Therefore, we will show a weaker statement, namely that for each ϵ greater than zero it holds true that

$$\limsup_{n\to\infty} \mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] \leq 1 + \epsilon.$$

Let $p_{\text{color}} : \mathbb{N} \to [0,1]$ denote the probability mass function that determines F_{color}. Without loss of generality we assume that $p_{\text{color}}(i) \geq p_{\text{color}}(i+1)$. Let $\gamma > 0$ be a small constant depending on ϵ. We denote by $\mathcal{J}_\gamma$ the smallest integer such that $\sum_{j=1}^{\mathcal{J}_\gamma} p_{\text{color}}(j) \geq 1-\gamma$. We call the colors $\{1, \dots, \mathcal{J}_\gamma\}$ *large*, and the remaining ones *small*. In our analysis we consider the number of opened bins by A for items of large colors and small colors separately. Therefore, we will make the following assumption:

▶ **Assumption 4.** *For all $j \in \mathbb{N}$ we have $p_{\text{color}}(j) > 0$.*

Otherwise, if the set of colors c with $p_{\text{color}}(c) > 0$ is finite, we can treat all these colors as large colors.

Now, the analysis is similar to the ones seen before: Let m be the smallest integer such that $2^{m+1} \geq n$ and $t = \lceil m/2 \rceil + 1$. The number of opened bins for the items in the first phases $P_0, \dots, P_{t-1}$ can be bounded from above by $8\sqrt{n}$. For $i \in \{t, \dots, m-1\}$ we have $|P_i| \geq \sqrt{n}$. Moreover, we assume without loss of generality that also $|P_m| \geq \sqrt{n}$.

Now, we analyze the number of opened bins more thoroughly. At first we consider the bins opened by the auxiliary algorithm. The number of opened bins for virtual items equipped with a small color can be controlled using Hoeffding's inequality two times.

▶ **Lemma 57.** *The number of bins opened for virtuals items equipped with a small color by the auxiliary algorithm is bounded from above by $\gamma|P_i| + 2\sqrt{|P_i^*|}\log(|P_i^*|)^{3/4}$ with probability at least $1 - 2\exp\left(-2\log(|P_i|)^{3/2}\right)$.*

It is more challenging to bound the number of opened bins for virtual items equipped with a large color:

▶ **Lemma 58.** *There exist universal constants α, $C_{\gamma,F}$ and $K_{\gamma,F}$ such that the number of opened bins by the the auxiliary algorithm for virtual items equipped with a large color is (for $|P_i|$ large enough) bounded from above by*

$$\frac{|P_i|}{|P_i^*|} \cdot \text{OPT}(P_i^*) + K_{\gamma,F}\sqrt{|P_i|}\log(|P_i|)^{3/4}$$

with probability at least $1 - C_{\gamma,F}\exp\left(-\alpha\log(|P_i|)^{3/2}\right)$.

Now, we turn to the analysis of the replacement of virtual items by real items: A simple application of Hoeffding's inequality makes it possible to control the number of real items of small colors that cannot replace a virtual item:

▶ **Lemma 59.** *The total number of reals items equipped with a small color in phase P_i that cannot replace a virtual item is bounded from above by $\gamma|P_i| + \sqrt{|P_i|}\log(|P_i|)^{3/4}$ with probability at least $1 - 2\exp\left(-2\log(|P_i|)^{3/2}\right)$.*

Again, the interesting part is to analyze the number of opened bins for real items equipped with large colors.

▶ **Lemma 60.** *There exist constants $\alpha_{\gamma,F}$, $C_{\gamma,F}$, and $K_{\gamma,F}$ (depending on γ and F) such that for $|P_i|$ large enough the number of real items, equipped with a large color, that cannot replace a virtual item is bounded from above by $K_{\gamma,F}\sqrt{|P_i|}\log(|P_i|)^{3/4}$ with probability at least $1 - C_{\gamma,F}\exp\left(-\alpha_{\gamma,F}\log(|P_i|)^{3/2}\right)$.*

Finally, we relate $\text{OPT}(P_i^*)$ with the value of an optimum solution of the whole instance:

▶ **Lemma 61.** *Let γ be greater than zero. Then, if n is large enough, it holds true for each P_i^* with $|P_i^*| \geq \sqrt{n}$ that*

$$\mathrm{OPT}(P_i^*) \leq \frac{|P_i^*|}{n} \cdot \mathrm{OPT}(I_n^F) + \gamma |P_i^*|$$

with probability at least $1 - 4\exp\left(-2\log(|P_i^*|)^{3/2}\right)$.

Combining Lemma 57, 58, 59, 60 and 61 we see that the number of opened bins in phases $P_t, \ldots, P_m$ is bounded from above by

$$\mathrm{OPT}(I_n^F) + 3\gamma n + K_{\gamma,F}\sqrt{n}\log(n)^{3/4} \tag{12}$$

with probability at least $1 - C_{\gamma,F}\exp\left(-\alpha_{\gamma,F}\log(n)^{3/2}\right)$. Here, constants may again differ from line to line.

Moreover, let S denote the size of a random item drawn according to F and let $\nu_F := \mathbb{E}[S]$. Then, using again Hoeffding's inequality we see that $S(I_n^F) \geq n \cdot \nu_F/2$ with high probability. Combining this with the estimates for the first phases, the estimate given in (12), and that the number of opened bins in the worst case is bounded from above by $2n$ we obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{E}\left[\frac{A(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] \\
\leq &\limsup_{n\to\infty}\left(1 + 6\nu_F\cdot\gamma + K_{\gamma,F}\frac{\log(n)^{3/4}}{\sqrt{n}} + 2n\cdot C_{\gamma,F}\exp\left(-\alpha_{\gamma,F}\log(n)^{3/2}\right)\right) \\
= &1 + 6\nu_F\cdot\gamma.
\end{aligned}$$

Since we can choose γ arbitrary small the proof of Proposition 36 follows.

3.2.4.2 Proof of Lemma 56

Here we will prove:

▶ **Lemma 56.** *Let F be a distribution, and let S be the size of a random item drawn according to F. If $\mathbb{P}[S > 0] > 0$ then it holds true that*

$$\limsup_{n\to\infty} \mathbb{E}\left[\frac{A^0(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] = 0.$$

Proof of Lemma 56. We show that for arbitrary δ greater than zero it holds

$$\limsup_{n\to\infty} \mathbb{E}\left[\frac{A^0(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] \leq \delta.$$

Let $\kappa > 0$ be arbitrary, and let u be such that $\sum_{j=1}^{u} p_{\mathrm{color}}(j) \geq 1 - \kappa$. Let $N(I_n^F)$ denote the number of items in I_n^F with colors in $\{u+1, u+2, \ldots\}$, S the size of a random drawn item according to F, and $\nu := \mathbb{E}[S]$.

Applying two times Hoeffding's inequality we obtain

$$\mathbb{P}\left[N(I_n^F) \geq 2\kappa n\right] \leq \exp\left(-2\kappa^2 n\right)$$

and

$$\mathbb{P}\left[S(I_n^F) < \frac{1}{2}\nu n\right] \leq \exp\left(-\frac{1}{2}\nu^2 n\right).$$

The number of opened bins for items of size zero can be bounded from above by $u + N(I_n^F)$. Therefore, it follows that

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{E}\left[\frac{A^0(I_n^F)}{\mathrm{OPT}(I_n^F)}\right] &\leq \limsup_{n\to\infty} \mathbb{E}\left[\frac{A^0(I_n^F)}{\max\{1, S(I_n^F)\}}\right] \\ &\leq \limsup_{n\to\infty}\left[\frac{u+2\kappa n}{\max\{1, \frac{1}{2}\nu n\}} + \left(\exp\left(-\frac{1}{2}\nu^2 n\right) + \exp\left(-2\kappa^2 n\right)\right) n\right] \\ &= \frac{4\kappa}{\nu}.\end{aligned}$$

Since we can choose κ arbitrary small, the statement follows. ◀

3.2.4.3 Proof of Lemma 57

Here we will prove:

▶ **Lemma 57.** *The number of bins opened for virtuals items equipped with a small color by the auxiliary algorithm is bounded from above by* $\gamma|P_i| + 2\sqrt{|P_i^*|}\log(|P_i^*|)^{3/4}$ *with probability at least* $1 - 2\exp\left(-2\log(|P_i|)^{3/2}\right)$.

Proof of Lemma 57. We estimate the number of opened bins by the auxiliary algorithm for virtual items equipped with small colors by the number $N_{i,\text{small}}^v$ of *all* generated virtual items of small colors.

Let $N_{i,\text{small}}^*$ denote the number of items of small colors in P_i^*. Then, virtual items with small colors are generated with probability $\hat{p}_{\text{small}} = N_{i,\text{small}}^*/|P_i^*|$. Applying Hoeffding's inequality we obtain

$$\mathbb{P}\left[N_{i,\text{small}}^v \leq |P_i| \cdot \hat{p}_{\text{small}} + \sqrt{|P_i|}\log(|P_i|)^{3/4} \mid N_{i,\text{small}}^*\right] \geq 1 - \exp\left(-2\log(|P_i|)^{3/2}\right).$$

Applying Hoeffding's inequality a second time it follows that

$$\mathbb{P}\left[N_{i,\text{small}}^* \leq |P_i^*| \cdot p_{\text{small}} + \sqrt{|P_i^*|}\log(|P_i^*|)^{3/4}\right] \geq 1 - \exp\left(-2\log(|P_i^*|)^{3/2}\right).$$

Combining both concentration bounds and using the fact that p_{small} is bounded from above by γ we see that $N_{i,\text{small}}^v$ is bounded from above by $\gamma|P_i| + 2\sqrt{|P_i^*|}\log(|P_i^*|)^{3/4}$ with probability at least $1 - 2\exp\left(-2\log(|P_i|)^{3/2}\right)$. ◀

3.2.4.4 Proof of Lemma 58

Here we will prove

▶ **Lemma 58.** *There exist universal constants* α, $C_{\gamma,F}$ *and* $K_{\gamma,F}$ *such that the number of opened bins by the the auxiliary algorithm for virtual items equipped with a large color is (for* $|P_i|$ *large enough) bounded from above by*

$$\frac{|P_i|}{|P_i^*|} \cdot \mathrm{OPT}(P_i^*) + K_{\gamma,F}\sqrt{|P_i|}\log(|P_i|)^{3/4}$$

with probability at least $1 - C_{\gamma,F}\exp\left(-\alpha\log(|P_i|)^{3/2}\right)$.

Proof of Lemma 58. There are two possible reasons that cause the auxiliary algorithm A' to open a new bin for a virtual item V of a large color:

- V has rank 1 in the model-packing.
- V cannot replace a second-order virtual item.

We start by estimating the number of opened bins because of virtual items of large colors with rank 1. We bound this term from above by the number of drawn items of rank 1 with *arbitrary* colors. Among the $|P_i^*|$ many items there are $\mathrm{OPT}(P_i^*)$ many items of rank 1. Applying Hoeffding's inequality we obtain that the number of opened bins because of an item with rank 1 is bounded from above by

$$\frac{\mathrm{OPT}(P_i^*)}{|P_i^*|} \cdot |P_i| + \sqrt{|P_i|}\log(|P_i|)^{3/4}$$

with probability at least $1 - \exp\left(-2\log(|P_i|)^{3/2}\right)$.

Now we want to bound the number of opened bins for virtual items that cannot replace second-order virtual items. Let c be a fixed large color. In the corresponding analysis of the auxiliary algorithm for cardinality-constrained bin packing we were able to use a union bound since the number of ranks was bounded from above by the parameter k. Here, the analysis is a bit more complicated as we pack all virtual items of a color c that cannot replace a second-order virtual item using Next-Fit. Therefore, we will bound the *total weight* T_c of these items. In order to do this, we proceed as in [72].

As before at the beginning of a new phase a model-packing $\mathcal{M}(P_i^*)$ of all items in P_i^* is computed. Let $N_{c,u}$ denote the number of virtual items with color c that cannot replace a second-order virtual item and for which the corresponding item in the model-packing has rank u. Moreover, let $m_{c,u}$ be the number of items of color c and rank u in $\mathcal{M}(P_i^*)$.

Let u be a fixed rank. We observe that $N_{c,u}$ can be seen as the result of an instance of (M4) with a random instance size $L_{c,u}$.

We consider two cases: If $L_{c,u} \leq |P_i|^{1/3}$, then we bound $N_{c,u}$ from above by $|P_i|^{1/3}$. So, now assume that $L_{c,u} > |P_i|^{1/3}$. Applying Lemma 27 we see that

$$\begin{aligned}
\mathbb{P}\left[N_{c,u} \leq K\sqrt{L_{c,u}}\log(L_{c,u})^{3/4} \mid L_{c,u}\right] &\geq 1 - C\exp\left(-\alpha\log(L_{c,u})^{3/2}\right)\\
&\geq 1 - C\exp\left(-\alpha\log(|P_i|^{1/3})^{3/2}\right)\\
&\geq 1 - C\exp\left(-\alpha\log(|P_i|)^{3/2}\right).
\end{aligned}$$

We still maintain the convention that constants may differ from line to line, but are always universal. Moreover, it follows from Hoeffding's inequality that we have

$$\mathbb{P}\left[L_{c,u} \leq \frac{2m_{c,u}}{|P_i^*|} \cdot |P_i| + \sqrt{|P_i|}\log(|P_i|)^{3/4}\right] \geq 1 - \exp\left(-2\log(|P_i|)^{3/2}\right).$$

Hence, using the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ and a union bound we see that

$$\begin{aligned}
N_{c,u} &\leq K\sqrt{L_{c,u}}\log(L_{c,u})^{3/4}\\
&\leq K\sqrt{L_{c,u}}\log(|P_i|)^{3/4}\\
&\leq K\sqrt{\frac{|P_i|}{|P_i^*|}}\log(|P_i|)^{3/4} \cdot \sqrt{m_{c,u}} + |P_i|^{1/4}\log(|P_i|)^{9/8}
\end{aligned}$$

with probability at least $1 - C\exp\left(-\alpha\log(|P_i|)^{3/2}\right)$, where α, C, and K are universal constants. Since we can draw virtual items of at most $|P_i|$ many different ranks, we can apply a union bound and see that this inequality holds for all ranks u with $L_{c,u} > 0$ simultaneously with probability at least

$$1 - |P_i| C \exp\left(-\alpha \log(|P_i|)^{3/2}\right) \geq 1 - C \exp\left(-\alpha \log(|P_i|)^{3/2}\right)$$

for an adjusted choice of α if $|P_i|$ is large enough.

Applying the estimates for both cases we obtain an upper bound $H_{c,u}$ with

$$N_{c,u} \leq H_{c,u} := |P_i|^{1/3} + K\sqrt{\frac{|P_i|}{|P_i^*|}} \cdot \log(|P_i|)^{3/4} \cdot \sqrt{m_{c,u}} + |P_i|^{1/4} \log(|P_i|)^{9/8}$$

with probability at least $1 - C \exp\left(-\alpha \log(|P_i|)^{3/2}\right)$.

Now we are able to give a bound for the total sum T_c of all unmatched items of color c. The size of items of rank u is bounded from above by $1/u$. Hence, we have $T_c \leq \sum_{u=2}^{|P_i^*|} N_{c,u}/u$. Let s_c be equal to $|P_i^*|$ if $\sum_{u=2}^{|P_i^*|} H_{c,u} \leq |P_i|$ and otherwise the smallest integer such that $\sum_{u=2}^{s_c} H_{c,u} \geq |P_i|$. Then, we have

$$\sum_{u=2}^{|P_i^*|} \frac{N_{c,u}}{u} \leq \sum_{u=2}^{s_c} \frac{H_{c,u}}{u}.$$

Since $H_{c,u} \geq |P_i|^{1/3}$ we have $s_c \leq |P_i|$, and therefore $\sum_{u=2}^{s_c} \frac{1}{u} \leq K \log(|P_i|)$. Applying Cauchy-Schwarz we obtain for $|P_i|$ large enough

$$\begin{aligned}
&\sum_{u=2}^{|P_i^*|} \frac{N_{c,u}}{u} \\
\leq &\sum_{u=2}^{s_c} \frac{1}{u} \left[|P_i|^{1/3} + K\sqrt{\frac{|P_i|}{|P_i^*|}} \cdot \log(|P_i|)^{3/4} \cdot \sqrt{m_{c,u}} + |P_i|^{1/4} \log(|P_i|)^{9/8} \right] \\
\leq &K|P_i|^{1/3} \log(|P_i|) + |P_i|^{1/4} \log(|P_i|)^{17/8} + K\sqrt{\frac{|P_i|}{|P_i^*|}} \cdot \log(|P_i|)^{3/4} \cdot \sum_{u=2}^{s_c} \frac{\sqrt{m_{c,u}}}{u} \\
\leq &K\sqrt{|P_i|} \log(|P_i|)^{3/4} + \log(|P_i|)^{3/4} \cdot \left(\sum_{u=2}^{s_c} m_{c,u} \right)^{1/2} \cdot \left(\sum_{u=1}^{\infty} \frac{1}{u^2} \right)^{1/2} \\
\leq &K\sqrt{|P_i|} \log(|P_i|)^{3/4}
\end{aligned}$$

with probability at least $1 - C \exp\left(-\alpha \log(|P_i|)^{3/2}\right)$ for universal constants α, C and K.

Now, applying a union bound on all large colors we see that there exist universal constants α, $C_{\gamma,F}$ and $K_{\gamma,F}$ such that the number of opened bins for virtual items of large items is bounded from above by $K_{\gamma,F}\sqrt{|P_i|} \log(|P_i|)^{3/4}$ with probability at least $1 - C_{\gamma,F} \exp\left(-\alpha \log(|P_i|)^{3/2}\right)$. ◀

3.2.4.5 Proof of Lemma 60

Here we will prove

▶ **Lemma 60.** *There exist constants $\alpha_{\gamma,F}$, $C_{\gamma,F}$, and $K_{\gamma,F}$ (depending on γ and F) such that for $|P_i|$ large enough the number of real items, equipped with a large color, that cannot replace a virtual item is bounded from above by $K_{\gamma,F}\sqrt{|P_i|} \log(|P_i|)^{3/4}$ with probability at least $1 - C_{\gamma,F} \exp\left(-\alpha_{\gamma,F} \log(|P_i|)^{3/2}\right)$.*

Proof of Lemma 60. We bound the number of real items that cannot replace a virtual item for each large color independently. Afterwards, the result follows by using a union bound.

When we compare the proof of this statement with the corresponding statement for cardinality-constrained bin packing (that is Lemma 39) we notice two obstacles:

- It is not guaranteed that the color of the generated virtual item is the same as of the corresponding real item.
- The distribution of the sizes of the virtual items with color c is only an approximation of the underlying distribution F_c.

The second point also occurs in the proof in [72] and will be solved using the Dvoretzky-Kiefer-Wolfowitz inequality.

Let c be a fixed large color, and we denote the probability of drawing an item of color c by p_c. At first we want to reduce the problem of matching real and virtual items of color c to an instance of matching variant (M3). Let $P_i = (A_{2^i+1}, \ldots, A_{\min\{2\cdot 2^i, n\}})$ be the sequence of real items and $(V_{2^i+1}, \ldots, V_{\min\{2\cdot 2^i, n\}})$ the corresponding virtual items. We denote by $\mathcal{J}_c^r = \{j_1^r, \ldots, j_a^r\}$ the indices of all *real* items in phase P_i such that $A_{j_m^r}$ is of color c. Moreover, we denote by $\mathcal{J}_c^v = \{j_1^v, \ldots, j_b^v\}$ the indices of all *virtual* items of color c.

We set $X_j := I_{\left\{A_{2^i+j} \text{ is of color } c\right\}}$, $Y_j := I_{\left\{V_{2^i+j} \text{ is of color } c\right\}}$ and $D^i(m) := \sum_{j=1}^{m} (X_j - Y_j)$. Now, let

$$\tilde{\mathcal{J}}_c^r := \left\{ j_u^r \in \mathcal{J}_c^r \,:\, D^i(j_u^r) \leq \max_{0 \leq m < j_u^r} D^i(m) \right\} = \left\{\tilde{j}_1^r, \ldots, \tilde{j}_{\tilde{a}}^r\right\}.$$

Set $L := \min\left\{|\tilde{\mathcal{J}}_c^r|, |\mathcal{J}_c^v|\right\}$.

Now, $\mathcal{P}^+ = \{(2u, A_{\tilde{j}_u^r})\}_{1 \leq u \leq L}$ and $\mathcal{P}^- = \{(2u-1, V_{j_u^v})\}_{1 \leq u \leq L}$ is an instance for (M3), and this instance was constructed in a way such that every matching for this instance is also a feasible matching for the original problem. Here the distribution of the sizes of the real items is given by F_c and the distribution for the virtual items is given by the empirical distribution induced by P_i^*.

The next steps are as follows:

- At first we will show a lower bound for the instance size L, which holds true with high probability.
- Afterwards, we will control the distance between F_c and the empirical distribution by applying the Dvoretzky-Kiefer-Wolfowitz inequality.
- Then, the statement follows for a fixed color c by applying Lemma 26.

Let $N_{i,c}^r$ denote the number of real items of color c in phase P_i, $N_{i,c}^*$ the number of items of color c in P_i^* and $N_{i,c}^v$ the number of virtual items generated in phase P_i of color c. Then, a virtual item of color c is generated with probability $\hat{p}_c = N_{i,c}^* / |P_i^*|$.

Applying two times Hoeffding's inequality we obtain

$$\mathbb{P}\left[|N_{i,c}^r - |P_i| \cdot p_c| \leq \sqrt{|P_i|} \log(|P_i|)^{3/4}\right] \geq 1 - 2\exp\left(-2\log(|P_i|)^{3/2}\right), \tag{13}$$

and

$$\mathbb{P}\left[|N_{i,c}^* - |P_i^*| \cdot p_c| \leq \sqrt{|P_i^*|} \log(|P_i^*|)^{3/4}\right] \geq 1 - 2\exp\left(-2\log(|P_i^*|)^{3/2}\right). \tag{14}$$

Moreover, we have $\mathbb{E}\left[X_j - Y_j | \, N^*_{i,c}\right] = p_c - \hat{p}_c$. We set $Z_j := X_j - Y_j$, and $\tilde{Z}_j := Z_j - (p_c - \hat{p}_c)$. Applying Proposition 9 and afterwards Hoeffding's inequality with Z_j it follows that there exist universal constants such that

$$\max_{0 \leq m \leq |P_i|} D^i(m) = \max_{0 \leq m \leq |P_i|} \sum_{j=1}^{m} Z_j \leq |P_i| \cdot (p_c - \hat{p}_c) + K\sqrt{|P_i|}\log(|P_i|)^{3/4}$$

with probability at least $1 - C\exp\left(-\alpha \log(|P_i|)^{3/2}\right)$.

Assuming that $\left|N^*_{i,c} - |P^*_i| \cdot p_c\right| \leq \sqrt{|P^*_i|}\log(|P^*_i|)^{3/4}$ it follows that

$$\max_{0 \leq m \leq |P_i|} D^i(m) \leq |P_i| \cdot (p_c - \hat{p}_c) + K\sqrt{|P_i|}\log(|P_i|)^{3/4} \leq K\sqrt{|P^*_i|}\log(|P^*_i|)^{3/4}. \tag{15}$$

Hence, combining (13), (14), and (15) it follows that there exist universal constants α, C and K such that $|\tilde{\mathcal{J}}^r_c| \geq p_c|P_i| - K\sqrt{|P^*_i|}\log(|P^*_i|)^{3/4}$ with probability at least $1 - C\exp\left(-\alpha \log(|P_i|)^{3/2}\right)$.

Furthermore, we can show in the same way that

$$\mathbb{P}\left[N^v_{i,c} \geq |P_i| \cdot p_c - 2\sqrt{|P_i|}\log(|P_i|)^{3/4}\right] \geq 1 - 4\exp\left(-\log(|P_i|)^{3/2}\right).$$

Thus, it follows that there exist universal constants α, C and K such that the size L of our matching instance is at least $|P_i|p_c - K\sqrt{|P_i|}\log(|P_i|)^{3/2}$ with probability at least $1 - C\exp\left(-\alpha \log(|P_i|)^{3/2}\right)$.

The virtual items of color c are drawn according to an empirical distribution $\hat{F}_{i,c}$. Now we will bound the distance $d(F_c, \hat{F}_{i,c})$, described in Lemma 26, between both probability measures.

Applying the Dvoretzky-Kiefer-Wolfowitz inequality we obtain

$$\mathbb{P}\left[d(F_c, \hat{F}_{i,c}) \leq \frac{\log(|P_i|)^{3/4}}{\sqrt{|P_i|}} \mid N^*_{i,c}\right] \geq 1 - 2\exp\left(-2N^*_{i,c} \cdot \frac{\log(|P_i|)^{3/2}}{|P_i|}\right).$$

So, assuming that $N^*_{i,c} \geq |P^*_i|p_c - \sqrt{|P^*_i|}\log(|P_i|^*)^{3/4}$ it follows that we obtain for $|P_i|$ large enough

$$\mathbb{P}\left[d(F_c, \hat{F}_{i,c}) \leq \frac{\log(|P_i|)^{3/4}}{\sqrt{|P_i|}} \mid N^*_{i,c}\right] \geq 1 - 2\exp\left(-p_c \log(|P_i|)^{3/2}\right).$$

Combining the previous concentration bounds for L and $d(F_c, \hat{F}_{i,c})$ and applying Lemma 26 it follows that there exist universal constants α_c, C and K such that the number of real items of color c that cannot replace a virtual item is bounded from above by $K\sqrt{|P_i|}\log(|P_i|)^{3/4}$ with probability at least $1 - C\exp\left(-\alpha_c \log(|P_i|)^{3/2}\right)$.

Hence, applying a union bound on all large colors the statement follows. ◀

3.2.4.6 Proof of Lemma 61

Here we will prove

▶ **Lemma 61.** *Let γ be greater than zero. Then, if n is large enough, it holds true for each P^*_i with $|P^*_i| \geq \sqrt{n}$ that*

$$\mathrm{OPT}(P^*_i) \leq \frac{|P^*_i|}{n} \cdot \mathrm{OPT}(I^F_n) + \gamma|P^*_i|$$

with probability at least $1 - 4\exp\left(-2\log(|P^*_i|)^{3/2}\right)$.

Proof of Lemma 61. The estimate we want to show is equivalent to

$$\frac{1}{|P_i^*|}\operatorname{OPT}(P_i^*) \leq \frac{1}{n}\operatorname{OPT}(I_n^F) + \gamma.$$

It holds

$$\begin{aligned}&\left|\frac{1}{|P_i^*|}\operatorname{OPT}(P_i^*) - \frac{1}{n}\operatorname{OPT}(I_n^F)\right| \\ &\leq \left|\frac{1}{|P_i^*|}\left(\operatorname{OPT}(P_i^*) - \mathbb{E}\left[\operatorname{OPT}(P_i^*)\right]\right)\right| \\ &\qquad + \left|\frac{1}{n}\left(\operatorname{OPT}(I_n^F) - \mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]\right)\right| + \left|\frac{1}{|P_i^*|}\mathbb{E}\left[\operatorname{OPT}(P_i^*)\right] - \frac{1}{n}\mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]\right|.\end{aligned}$$

Applying two times Hoeffding's inequality it follows that

$$\begin{aligned}\mathbb{P}\left[\frac{1}{|P_i^*|}\left|\operatorname{OPT}(P_i^*) - \mathbb{E}\left[\operatorname{OPT}(P_i^*)\right]\right| \geq \frac{\log(|P_i^*|)^{3/4}}{\sqrt{|P_i^*|}}\right] &\leq 2\exp\left(-2\log(|P_i^*|)^{3/2}\right) \\ \mathbb{P}\left[\frac{1}{n}\left|\operatorname{OPT}(I_n^F) - \mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]\right| \geq \frac{\log(n)^{3/4}}{\sqrt{n}}\right] &\leq 2\exp\left(-2\log(n)^{3/2}\right).\end{aligned}$$

Moreover, it holds that $\mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]/n \leq 1$, and since OPT is a subadditive function it follows from Fekete's Lemma (see [40]) that $\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]$ exists. Thus, it follows that for n large enough we have

$$\left|\frac{1}{|P_i^*|}\mathbb{E}\left[\operatorname{OPT}(P_i^*)\right] - \frac{1}{n}\mathbb{E}\left[\operatorname{OPT}(I_n^F)\right]\right| \leq \frac{\gamma}{3}.$$

Furthermore, for n large enough we have also $\log(|P_i^*|)^{3/4}/\sqrt{|P_i^*|} \leq \gamma/3$. So, all in all we have for n large enough

$$\operatorname{OPT}(P_i^*) \leq \frac{|P_i^*|}{n}\operatorname{OPT}(I_n^F) + \gamma|P_i^*|$$

with probability at least $1 - 4\exp\left(-2\log(|P_i^*|)^{3/2}\right)$. ◀

3.2.5 A Lower Bound for Class-constrained Bin Packing with respect to the Random-order Ratio

Here, we will prove:

▶ **Proposition 37.** *Consider online class-constrained bin packing with parameter k equal to 2. Then, for all deterministic online algorithms A it holds that* $\operatorname{RR}(A) \geq {}^{10}/_{9}$.

Proof of Proposition 37. Let $a \in \mathbb{N}$, and let $\gamma \leq {}^{1}/_{2a}$. The set of items $\mathcal{I}_a$ we will deal with is the following: For each of the colors in $[2a]$, there are n items of size γ/n. Furthermore, there are a items of size $1-2\gamma$. We call these items *special* and will specify their colors later. So in total there $2a \cdot n + a$ many items in $\mathcal{I}_a$.

Let Π denote the set of all possible permutations of $[2an + a]$. For $\sigma \in \Pi$ let I_a^σ denote the instance where the items arrive in the order specified by σ. We denote by $\mathcal{E}$ the event that for each color at least one item of size γ/n arrives before the first special item is drawn.

▶ **Lemma 62.** *Let $\epsilon > 0$ be arbitrary. For sufficiently large n, we have $\mathbb{P}[\mathcal{E}] \geq 1 - \epsilon$.*

Proof. For $i \in [2a]$ let T_i denote the time point, when the first item of size γ/n with color i arrives. Moreover, let T_s denote the time point, when the first special items arrives. As there are n items of size γ/n with color i and a special items, we have $\mathbb{P}[T_s < T_i] = \frac{a}{a+n}$. Using a union bound, it follows that.

$$\mathbb{P}[\mathcal{E}] \geq 1 - 2a \cdot \frac{a}{a+n} = 1 - \frac{2a^2}{a+n}.$$

So, choosing n large enough, we obtain the statement. ◀

Now we consider an arbitrary but fixed online algorithm A for class-constrained bin packing. For each $\sigma \in \Pi$ let t_σ denote the point in time when the first special item arrives. For $1 \leq i < j \leq 2a$ we denote by $m_{ij}(\sigma)$ the number of bins opened by A at t_σ that contain at least one item of color i and at least one item of color j when I_a^σ is processed. Moreover, for $1 \leq i \leq 2a$, $m_i(\sigma)$ denotes the number of opened bins at t_σ that contain only items of color i.

If $\sigma \in \mathcal{E}$, we have

$$2 \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma) + \sum_{1 \leq i \leq 2a} m_i(\sigma) \geq 2a,$$

since in this case $2a$ colors were drawn. Therefore, it is

$$\begin{aligned} \sum_{\sigma \in \Pi} A(I_a^\sigma) &\geq \sum_{\sigma \in \Pi} \sum_{1 \leq i \leq 2a} m_i(\sigma) \\ &\geq \sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i \leq 2a} m_i(\sigma) \geq 2a|\mathcal{E}| - 2 \sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma). \end{aligned} \tag{16}$$

This is the first lower bound we obtain for the number of bins the algorithm has to open. It basically states that we have to open many bins that contain two different colors, otherwise the performance of the algorithm cannot be better than close to 2.

Now, we want to show that if we open many bins that contain two different colors, we can find a mapping of the special items to the colors, such that the algorithm opens a constant fraction of bins that are nearly empty in the end (that is, they contain only items of size γ/n).

▶ **Lemma 63.** *There exists a subset $\{j_1, \dots, j_a\} = \mathcal{J}$ of $[2a]$, such that*

$$\sum_{\sigma \in \mathcal{E}} \sum_{j_1, j_2 \in \mathcal{J} : j_1 \neq j_2} m_{j_1 j_2}(\sigma) \geq \frac{a-1}{4a-2} \sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma). \tag{17}$$

Proof. The statement can be shown using the probabilistic method. Assume that the set $\mathcal{J}$ is generated by drawing a indices from $[2a]$ using sampling without replacement. Then, we obtain

$$\begin{aligned} \mathbb{E}\left[\sum_{\sigma \in \mathcal{E}} \sum_{j_1, j_2 \in \mathcal{J} : j_1 \neq j_2} m_{j_1 j_2}(\sigma)\right] &= \sum_{1 \leq i < j \leq 2a} \mathbb{P}[i, j \in \mathcal{J}] \sum_{\sigma \in \mathcal{E}} m_{ij}(\sigma) \\ &= \frac{a-1}{4a-2} \sum_{1 \leq i < j \leq 2a} \sum_{\sigma \in \mathcal{E}} m_{ij}(\sigma). \end{aligned}$$

Therefore, there must exist a subset $\mathcal{J}$ that fulfills the requirements. ◀

Let $\{j_1, \ldots, j_a\} = \mathcal{J}$ denote a subset of $[2a]$ that satisfies condition (17). We set $\{j'_1, \ldots, j'_a\} = \mathcal{J}' = [2a] \backslash \mathcal{J}$, and call the colors in $\mathcal{J}$ *small* and the colors in $\mathcal{J}'$ *large*. Then, each special item will be equipped with a different large color.

We can pack the items in $\mathcal{I}_a$ in a way that all items of a single large color and all items of a single small color are packed into one bin. Hence, we have $\mathrm{OPT}(\mathcal{I}_a) = a$.

We observe that A cannot place an item of a large color into a bin that already contains two small colors. Therefore, we obtain

$$\sum_{\sigma \in \Pi} A(I_a^\sigma) \geq \frac{a-1}{4a-2} \sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma) + a|\mathcal{E}|. \tag{18}$$

This is our second lower bound for the number of bins opened by A.

Now, combining (16) and (18) it follows that

$$\frac{\mathbb{E}\left[A(I_a^\sigma)\right]}{\mathrm{OPT}(I_a)} \geq \frac{1}{a|\Pi|} \max \left\{ 2a|\mathcal{E}| - 2 \sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma), \frac{a-1}{4a-2} \sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma) + a|\mathcal{E}| \right\}.$$

The maximum is minimized if we set $\sum_{\sigma \in \mathcal{E}} \sum_{1 \leq i < j \leq 2a} m_{ij}(\sigma) = \frac{4a-2}{9a-5} \cdot a|\mathcal{E}|$.

Then, using Lemma 62, we obtain the estimate

$$\frac{\mathbb{E}\left[A(I_a^\sigma)\right]}{\mathrm{OPT}(\mathcal{I}_a)} \geq \left(1 + \frac{a-1}{9a-5}\right) \cdot (1 - \epsilon).$$

This yields the proof of the statement. ◀

4 Breaking Worst-case Bounds in Bounded-space Bin Packing and Covering in the Random-order Model

In the previous section we have discussed the complexity of classical, cardinality-constrained and class-constrained bin packing with respect to probabilistic performance measures.

We have seen that, apart from class-constrained bin packing in the random-order model, it is possible to construct algorithms that behave very well. But we have also seen that those algorithms are only of theoretical interest as their runtime is very high. Moreover, in practical applications it is often reasonable to assume that algorithms can only use a restricted number of open bins simultaneously. That is, they are bounded-space algorithms. We call an algorithm that uses only K bins at the same time a *K-bounded space algorithm.*

In this section we show that there exist simple heuristics that benefit from items arriving in random order, that is it is possible to obtain a random-order ratio that is better than the corresponding bound for the competitive ratio. The most important results of this section are compared to existing ones in Figure 14, 15 and 16.

	lower bound	upper bound
CR(A)	≈ 1.691 [58]	$1.691 + \epsilon$ [58]
RR(A)	$1 + \frac{1}{4K+4}$ [25]	?
APR($A, \mathcal{D}$)	$1 + \frac{1}{4K+4}$ [25]	?

	lower bound	upper bound
CR(A)	≈ 1.691 [58]	$1.691 + \epsilon$ [58]
RR(A)	$1 + \frac{1}{4K+4}$ [25]	**1.671**
APR($A, \mathcal{D}$)	$1 + \frac{1}{4K+4}$ [25]	**1.671**

Figure 14 Comparison between previously known bounds that are universally valid for bounded-space algorithms for classical bin packing on the left hand side and our results on the right hand side.

	lower bound	upper bound
CR(NF)	2 [50]	2 [50]
RR(NF)	2 [20]	2
APR(NF, $\mathcal{D}$)	4/3 [23]	?
CR(SNF)	2 [69]	2 [69]
RR(SNF)	?	?
APR(SNF, $\mathcal{D}$)	1.227 [69]	?
CR(WF)	2 [50]	2 [50]
RR(WF)	?	?
APR(WF, $\mathcal{D}$)	?	?

	lower bound	upper bound
CR(NF)	2 [50]	2 [50]
RR(NF)	2 [20]	2
APR(NF, $\mathcal{D}$)	**2**	**2**
CR(SNF)	2 [69]	2 [69]
RR(SNF)	**2**	**2**
APR(SNF, $\mathcal{D}$)	**2**	**2**
CR(WF)	2 [50]	2 [50]
RR(WF)	**2**	**2**
APR(WF, $\mathcal{D}$)	**2**	**2**

Figure 15 Comparison between previously known bounds for specific algorithms for classical bin packing on the left hand side and our results on the right hand side.

4.1 Bounded-Space Online Bin Packing

Before showing the existence of a 4-bounded space algorithm A with RR(A) < 1.671 we take a closer look at the result presented in [20]. In this publication the authors show that the simple 1-bounded-space heuristic Next-Fit obtains a random-order ratio of 2. That is, this heuristic does not benefit from a randomized order of the input compared to the worst case. As Next-Fit is the only reasonable 1-bounded space heuristic, this also implies that at least two bins are needed to beat the lower bound for the competitive ratio of bounded-space algorithm $h_\infty \approx 1.691$ given in [58]. We want to take up the statement regarding the

	lower bound	upper bound		lower bound	upper bound
CR(A)	$1/2$ [4, 5]	$1/2$ [28]	CR(A)	$1/2$ [4, 5]	$1/2$ [28]
CR(DNF)	$1/2$ [4, 5]	$1/2$ [4, 5]	CR(DNF)	$1/2$ [4, 5]	$1/2$ [4, 5]
RR(DNF)	?	0.736 [24]	RR(DNF)	**0.502**	$\mathbf{0.\bar{6}}$
APR(DNF, $\mathcal{D}$)	?	0.736 [24]	APR(DNF, $\mathcal{D}$)	**0.502**	$\mathbf{0.\bar{6}}$
CR(DH$_K$)	$1/2$ [17]	$1/2$ [17]	CR(DH$_K$)	$1/2$ [17]	$1/2$ [17]
RR(DH$_K$)	$1/2$ [17]	$1/2$ [17]	RR(DH$_K$)	$1/2$ [17]	$1/2$ [17]
APR(DH$_K$, $\mathcal{D}$)	?	0.71 [17]	APR(DH$_K$, $\mathcal{D}$)	**0.5**	**0.5**

Figure 16 Comparison between previously known bounds for algorithms for classical bin covering on the left hand side and our results on the right hand side.

random-order ratio of Next-Fit and generalize it in several ways: The following proposition shows that the statement is not only true for Next-Fit in the random-order model, but also for the heuristics Smart-Next-Fit and Worst-Fit with respect to the average performance ratio. Moreover, we will investigate the parametric case, where a parameter $k \in \mathbb{N}$ with $k \geq 2$ is given, and all item sizes are bounded from above by $1/k$. The parametric case is studied for example in [10, 50, 82].

▶ **Proposition 64.** *Let $A \in \{\mathrm{NF}, \mathrm{SNF}, \mathrm{WF}\}$ and $k \in \mathbb{N}$ with $k \geq 2$. Let $\mathcal{D}_k$ denote the set of all distributions, where the maximum item size is bounded from above by $1/k$. Then, we have* $\mathrm{APR}(A, \mathcal{D}_k) = 1 + \frac{1}{k-1}$.

▶ Remark. It follows that the same is also true for the random-order ratio restricted to instances with maximal item size bounded by $1/k$. Furthermore, the competitive ratio of the algorithms Next-Fit, Smart-Next-Fit, and Worst-Fit in the case that the item sizes are bounded from above by $1/k$ is also $1 + 1/(k-1)$. So, the algorithms also do not benefit from randomization if the item sizes are much smaller than 1.

Proof. Since Next-Fit, Smart-Next-Fit and Worst-Fit are 2-competitive algorithms, we can apply Lemma 31. Therefore, it suffices to show a lower bound of $\frac{k}{k-1} - \epsilon$ for $\frac{\mathbb{E}[A(I_n^F)]}{\mathbb{E}[\mathrm{OPT}(I_n^F)]}$ for arbitrary ϵ.

Let F be given by $\mathcal{I} = \{\frac{1}{k}, \frac{1}{a^2}\}$ with $p(\frac{1}{k}) = \frac{k}{a+k}$ and $p(\frac{1}{a^2}) = \frac{a}{a+k}$, where $a \in \mathbb{N}$ with $a > \sqrt{k}$. We can think of this distribution as follows: There are a bins each containing k items of size $1/k$ and one additional bin containing a^2 items of size $1/a^2$. The distribution is then given by the uniform distribution on the set of these items. Items of size $1/k$ will be called *big* and items of size $1/a^2$ *small*.

We notice that all three algorithms share the same property: All opened bins – possibly apart from the last – have bin level greater than $1 - 1/k$.

So, if the predecessor of a big item is a small item it is not possible to put this item in a bin already containing $k-1$ many big items. Such big items will be called *nice*. Let $N(I_n^F)$ denote the number of nice items in the input. It follows that A needs to open at least $N(I_n^F)/(k-1)$ many bins.

Hence, we have

$$\begin{aligned}\mathbb{E}\left[A(I_n^F)\right] &\geq \frac{1}{k-1} \cdot \mathbb{E}\left[N(I_n^F)\right] \\ &= \frac{1}{k-1} \cdot (n-1) \cdot \frac{a}{a+k} \cdot \frac{k}{a+k} \\ &= (n-1) \cdot \frac{k}{k-1} \cdot \frac{a}{(a+k)^2}.\end{aligned}$$

Furthermore, we can give an upper bound for $\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]$ as follows: Let $b(I_n^F)$ denote the number of big items, and $sm(I_n^F)$ the number of small items in I_n^F. We can pack the small and big items separately using Next-Fit for each type of item. Then, we have

$$\begin{aligned}\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] &\leq 2 + \frac{1}{k} \cdot \mathbb{E}\left[b(I_n^F)\right] + \frac{1}{a^2} \cdot \mathbb{E}\left[sm(I_n^F)\right] \\ &= 2 + \frac{1}{k} \cdot n \cdot \frac{k}{a+k} + \frac{1}{a^2} \cdot n \cdot \frac{a}{a+k} \\ &= 2 + n \cdot \frac{a+1}{a(a+k)}.\end{aligned}$$

It follows that for arbitrary $\epsilon > 0$ we can find an a such that for sufficiently large n we have

$$\frac{\mathbb{E}\left[A(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \geq \frac{k}{k-1} - \epsilon.$$

◀

Now, we show that there exists a 4-bounded-space algorithm with random-order ratio smaller than h_∞. Our approach is based on the work of Lee and Lee [58], who presented the lower bound of h_∞ and introduced the family of Harmonic-algorithms. The algorithm HARMONIC$_4$ proceeds in the following way: The interval $[0, 1]$ is subdivided into four intervals $I_1 = (1/2, 1]$, $I_2 = (1/3, 1/2]$, $I_3 = (1/3, 1/3]$, and $I_4 = [0, 1/4]$. Then, HARMONIC$_4$ uses four open bins B_1, B_2, B_3 and B_4 and packs items from subinterval I_j in next-fit-manner in bin B_j.

We notice that the number of opened bins for the items in interval I_1, I_2 and I_3 is independent of the order the arrival of the items, in contrast to the number of opened bins for items from I_4. We want to give a short idea why $\mathrm{RR}(\text{HARMONIC}_4) \geq 1.7$: Think of a bins containing the items $1/2 + \epsilon$, $1/3 + \epsilon$ and $1/6 - 2\epsilon$. Furthermore, there is one additional bin containing lots of items of size 3ϵ. Now, assume that these items arrive in random order. HARMONIC$_4$ has to open a bins for items from I_1, and $a/2$ many bins for items from I_2. Moreover, choosing ϵ appropriately depending on a we can show as in the proof of Proposition 64 that we need with high probability $a/5$ many bins for packing the items from I_4. It follows that HARMONIC$_4$ will open approximately $1.7a$ many bins with high probability, while the optimum value is $a + 1$.

The algorithm A we will consider here, proceeds similar to HARMONIC$_4$ with only one slight adjustment: If an item x with size in I_2 arrives, we check if the currently open bin B_1 already contains an item y from I_1. If this is true, and it is $x + y \leq 1$, then we pack x in bin B_1 and close this bin. Otherwise, we put it in bin B_2. Obviously, this algorithm has the same worst-case guarantee as HARMONIC$_4$: A never opens more bins than HARMONIC$_4$, but if all items with sizes in I_2 arrive before the items from I_1, we do not benefit from our adjustment. So, according to [58] the competitive ratio of A is bounded from above by $31/18 \approx 1.722$. But in the following we will see that we are able to use the space in B_1 if the items arrive in random order:

▶ **Proposition 65.** *It holds* $\mathrm{RR}(A) < 1.671$*, where A is the described* 4*-bounded-space algorithm.*

Before starting to analyze the algorithm A, we will summarize the ideas from [58]: Their analysis is based on *weight functions*, a popular tool in bin packing, which is also applied in for example [52, 58, 82]. Afterwards, we show how to transfer the idea to our probabilistic setup.

Let $\mathcal{I}$ be an adversarially chosen set of items that arrives in the – also adversarially chosen – order I. For $i \in \{1,2,3\}$ let n_i denote the number of items in $\mathcal{I}$ that lie in interval I_i. Furthermore, let s_4 denote the total size of all remaining items. Then, we can bound the number of opened bins of the algorithm as follows

$$\begin{aligned}\text{HARMONIC}_4(I) &\leq n_1 + \left\lceil \frac{1}{2}n_2 \right\rceil + \left\lceil \frac{1}{3}n_3 \right\rceil + \left\lceil \frac{4}{3}s_4 \right\rceil \\ &\leq n_1 + \frac{1}{2}n_2 + \frac{1}{3}n_3 + \frac{4}{3}s_4 + 3.\end{aligned}$$

Then, a *weight function* $g : \mathcal{I} \to [0,1]$ is defined. Here, the weight of an item roughly corresponds to the cost the algorithm incurs by packing this item, that is, the number of bins the algorithm has to open to pack the item. For one item from I_1, for two items from I_2 and for three items from I_3 the algorithm has to open a new bin. Furthermore, (apart from the last bin) every bin designated for items from I_4 has a bin level of at least $3/4$. Hence, for $x \in I_i$ with $i \in \{1,2,3\}$ we set $g(x) = 1/i$ and $g(x) = 4x/3$ for $x \in I_4$. Then, we have

$$\text{HARMONIC}_4(I) \leq \sum_{i=1}^{n} g(a_i) + 3.$$

Now, we fix an optimal packing $\mathcal{M}$ of the items in $\mathcal{I}$ and for $j \in \{1,\dots,\text{OPT}(\mathcal{I})\}$ let G_j denote the total weight of all items packed in the j-th bin of $\mathcal{M}$. Let

$$\mathcal{S} = \left\{ (y_1, y_2, \dots, y_t) \,:\, t \in \mathbb{N},\, 1 \leq i \leq t,\, y_i \geq 0,\, \sum_{i=1}^{t} y_i \leq 1 \right\}$$

denote the set of all possible partitions of 1. Finally, let $G^* = \sup_{(y_1,\dots,y_t)\in\mathcal{S}} \sum_{i=1}^{t} g(y_i)$ denote the maximum weight a bin can contain. Then, we have

$$\text{HARMONIC}_4(I) \leq \sum_{i=1}^{n} g(a_i) + 3 = \sum_{j=1}^{\text{OPT}(\mathcal{I})} G_j + 3 \leq G^* \cdot \text{OPT}(\mathcal{I}) + 3.$$

It follows that bounding the competitive ratio of HARMONIC_4 boils down to giving an upper bound for G^*.

In order to do so we note that $g(x) \leq 3x/2$ if $x \leq 1/2$, and $g(x) \leq 4x/3$ if $x \leq 1/3$. Now, we consider several cases: At first assume that a bin contains an item from I_1 and an item from I_2. Then, the remaining items in the bins come from I_4 and their total size is bounded from above by $1/6$. Hence, in this case the total weight of the bin is bounded from above by $1 + 1/2 + 1/6 \cdot 4/3 = 31/18 \approx 1.722$. In case that a bin contains one item from I_1, but no items from I_2 we have an upper bound for the weight of $1 + 1/2 \cdot 4/3 = 5/3$. Finally, if a bin contains no item from I_1, then the weight is bounded from above by $3/2$. Therefore, we see that HARMONIC_4 is a $31/18$-competitive algorithm.

Now, we want to analyze our proposed algorithm A in the random-order model: Let $\mathcal{I}$ be an adversarially chosen set of items that arrives in random order I^σ. By trying to pack items from I_2 in bin B_1, we reduce the number of items from I_2 that are packed in B_2 by $Q(I^\sigma)$. As previously defined for $i \in \{1,2,3\}$ let n_i denote the number of items from $\mathcal{I}$ that are in I_i and s_4 the total size of all remaining items. Then, we have

$$\mathbb{E}\left[A(I^\sigma)\right] \leq n_1 + \frac{1}{2}\left(n_2 - \mathbb{E}\left[Q(I^\sigma)\right]\right) + \frac{1}{3}n_3 + \frac{4}{3}s_4 + 3.$$

Again we fix an optimal packing $\mathcal{M}$ of the items in $\mathcal{I}$ and categorize the bins in $\mathcal{M}$ in three classes:

- Bins of the *first* type contain one item from I_1 and one item from I_2.
- Bins of the *second* type contain one item from I_1, but no item from I_2.
- Bins of the *third* type contain no item from I_1.

Let m_1, m_2, and m_3 denote the numbers of bins of the corresponding type in the fixed optimal solution.

We observe that $Q(I^\sigma)$ depends only on the items from I_1 and I_2. To give a lower bound for $\mathbb{E}[Q(I^\sigma)]$ we make without loss of generality the following two assumptions (A would only benefit from violating them):

▶ **Assumption 5.**
- *We assume that items, whose sizes are in I_2, and which are located in $\mathcal{M}$ in bins of the third type are so large, that they cannot be combined with any item from I_1.*
- *We assume that items, whose sizes are in I_1, and which are located in $\mathcal{M}$ in bins of the second type are so large, that they cannot be combined with any item from I_2.*

Then, we see that the items that are *relevant* for determining $\mathbb{E}[Q(I^\sigma)]$, are the items with sizes in I_1 or items with sizes in I_2 that are packed in $\mathcal{M}$ in a bin of the first type. Therefore, there are in total $2m_1 + m_2$ many relevant items.

Now, we assign the following weights $w : \mathcal{I} \to [0,1]$ to the items:

- If $x \in I_1$, then we set $w(x) = 1$.
- If $x \in I_3$, then we set $w(x) = 1/3$.
- If $x \in I_4$, then we set $w(x) = 4x/3$.
- If $x \in I_2$ and x is in $\mathcal{M}$ located in a bin of the first type, then we set $w(x) = 1/2 - \mathbb{E}[Q(I^\sigma)]/(2m_1)$.
- If $x \in I_2$ and x is in $\mathcal{M}$ not located in a bin of the first type, then we set $w(x) = 1/2$.

Now, we want to give a lower bound for $\mathbb{E}[Q(I^\sigma)]$. Let $(R_1, \ldots, R_{2m_1+m_2})$ be a random permutation of the relevant items. We will estimate the probability p that for two *consecutive* items the first one x is from I_1 and the second one y is from I_2 and we have $x + y \leq 1$:

▶ **Lemma 66.** *We have*

$$p \geq \frac{1}{2m_1 + m_2} \cdot \sum_{i=1}^{m_1} \frac{i}{2m_1 + m_2 - 1}.$$

Proof. As we assumed that all items from I_1 that are located in $\mathcal{M}$ in a bin of the second type are too large to be combined with an item from I_2, there are only m_1 relevant items from I_1 that can be combined with items from I_2. We call these items *nice*. Let $\ell_1 \geq \ldots \geq \ell_{m_1}$ denote the sizes of the nice items. Then, we observe that for the i-th nice item, there are at least i many items from I_2 such that they fit together in a bin. Hence, summing up yields

$$p \geq \frac{1}{2m_1 + m_2} \cdot \sum_{i=1}^{m_1} \frac{i}{2m_1 + m_2 - 1}.$$

◀

Then, using the previous lemma, we obtain

$$\begin{aligned}\mathbb{E}[Q(I^\sigma)] &\geq (2m_1 + m_2 - 1) \cdot p \\ &\geq (2m_1 + m_2 - 1) \cdot \frac{1}{2m_1 + m_2} \cdot \sum_{i=1}^{m_1} \frac{i}{2m_1 + m_2 - 1}\end{aligned}$$

$$\geq \frac{1}{2} \cdot \frac{m_1(m_1+1)}{2m_1+m_2}.$$

Proceeding as in the analysis of the worst case we see that the maximum weight W_1 of bins of the first type is bounded from above by $\frac{31}{18} - \frac{1}{4} \cdot \frac{m_1+1}{2m_1+m_2}$, the maximum weight W_2 of bins of the second type is bounded from above by $W_2 \leq 5/3$, and the maximum weight W_3 of bins of the third type is bounded from above by $3/2$. To give an upper bound for our algorithm, we have to choose m_1, m_2 and m_3 in such a way that we maximize

$$\frac{m_1 W_1 + m_2 W_2 + m_3 W_3}{m_1 + m_2 + m_3}.$$

Since p is independent of the bins of the third type, and we have $W_3 \leq W_2$ we can assume that $m_3 = 0$. Hence, it remains to maximize the expression $\frac{m_1 W_1 + m_2 W_2}{m_1+m_2}$. Let $m := m_1 + m_2$. Then, we have

$$\begin{aligned}
\frac{m_1 W_1 + m_2 W_2}{m_1 + m_2} &\leq \frac{m_1 \cdot \left(\frac{31}{18} - \frac{1}{4} \cdot \frac{m_1+1}{2m_1+m_2}\right) + \frac{5}{3} m_2}{m} \\
&\leq \max_{x\in[0,1]} \left[xm \cdot \left(\frac{31}{18} - \frac{1}{4} \cdot \frac{xm+1}{2xm+(1-x)m} \right) + \frac{5}{3}(1-x)m \right] \cdot \frac{1}{m} \\
&\leq \max_{x\in[0,1]} \left[\frac{5}{3} + \frac{31}{18}x - \frac{5}{3}x - \frac{1}{4}x \cdot \frac{xm+1}{xm+m} \right] \\
&\leq \frac{5}{3} + \max_{x\in[0,1]} \left[\frac{1}{18}x - \frac{1}{4} \cdot \frac{x^2}{x+1} \right].
\end{aligned}$$

Calculating the derivative we see that this term is maximized for $x = 3/\sqrt{7} - 1$. Then, we obtain an upper bound of $19/9 - \sqrt{7}/6 \approx 1.6702$. This shows Proposition 65.

The question arises if we can beat h_∞ using HARMONIC$_3$ with the mentioned modification. To investigate this question a more elaborate research is needed: If we consider again the worst case we find out that there exist bin configurations with total weight of 1.75 that do not contain an item from I_2. Hence, our algorithm would not benefit from the suggested modification. But, at first this bound is not necessarily tight and in the second place it is possible that the process of packing the items from I_3 benefits from the random arrival of the items. So, a more detailed analysis of packing items in B_3 would be needed to answer this question.

4.2 Bounded-Space Online Bin Covering

In the previous part about bounded-space online bin packing we have seen that the simple 1-bounded-space heuristic Next-Fit does not benefit from randomized input, that is the average performance ratio equals the competitive ratio – even in the parameterized case. For online bin covering it was shown that no online algorithm can obtain a competitive ratio better than $1/2$ and this bound is obtained by any reasonable algorithm, especially by Dual-Next-Fit [17, 28]. The main result of this part will be a lower bound for the random-order ratio of Dual-Next-Fit, which is (slightly) larger than $1/2$. So, Dual-Next-Fit behaves better than its counterpart for online bin packing and beats every algorithm in comparison with the competitive ratio.

4.2.1 Results

The main result is the following:

▶ **Theorem 67.** *It holds*

$$0.502 \leq \mathrm{RR}(\mathrm{DNF}) \leq \frac{2}{3}.$$

At this point we give a high-level overview of the proof by breaking it up into several lemmata. The proofs of the lemmata will be postponed to Section 4.2.2.

The investigation starts with observing that it suffices to analyze the behavior of the algorithm on a restricted set of instances:

▶ **Assumption 6.** *The investigated sets of items $\mathcal{I}$ satisfy* $\mathrm{OPT}(\mathcal{I}) = S(\mathcal{I})$*, where* $S(\mathcal{I})$ *denotes the total size of all items in* $\mathcal{I}$*.*

This assumption is justified by the observation that Dual-Next-Fit is a *monotone* algorithm: That is, if we obtain an instance I' from I by decreasing the sizes of items or deleting them, then we have $\mathrm{DNF}(I') \leq \mathrm{DNF}(I)$. We call instances with the property that $S(I) = \mathrm{OPT}(I)$ also *perfect-packing* instances.

In the following we will categorize the items into two types: Items with size greater than $1/2$ will be called *large* and otherwise they are called *small*. The following lemma shows that instances that contain only a small amount of large items are not critical for our analysis:

▶ **Lemma 68.** *Let I be an instance and let ℓ denote the number of large items in I. If* $\ell < 1/4 \cdot \mathrm{OPT}(I)$*, then we have* $\mathrm{DNF}(I) \geq \frac{7}{12}\,\mathrm{OPT}(I) - 2$*.*

Since we will only show an ϵ-improvement over the lower bound of $1/2$, this justifies the second assumption:

▶ **Assumption 7.** *If ℓ denotes the number of large items in $\mathcal{I}$, then we assume that* $\ell \geq 1/4\,\mathrm{OPT}(\mathcal{I})$*.*

We now turn to the analysis of the algorithm: For a random permutation I^σ of the items in $\mathcal{I}$ let $L_i(I^\sigma)$ denote the bin level of the i-th covered bin. In case that $i > \mathrm{DNF}(I^\sigma)$ we set $L_i(I^\sigma) = 0$. Furthermore, let $W(I^\sigma)$ denote the total size of all items in the last opened bin that is not covered. In case that such a bin does not exist we set $W(I^\sigma) = 0$. The *overshoot* of the i-th bin is defined as $R_i(I^\sigma) := L_i(I^\sigma) - 1$, and in case that $i > \mathrm{DNF}(I^\sigma)$ we set $R_i(I^\sigma) = 0$.

A key observation in the analysis is that the overshoot of the bins is *conditionally identically distributed*:

▶ **Lemma 69.** *Let $x \in [0,1]$ and $m \in \{1, \ldots, \mathrm{OPT}(\mathcal{I})\}$ with* $\mathbb{P}[\mathrm{DNF}(I^\sigma) = m] > 0$*. Then, for $i, j \leq m$ we have*

$$\mathbb{P}\left[R_i(I^\sigma) = x \mid \mathrm{DNF}(I^\sigma) = m\right] = \mathbb{P}\left[R_j(I^\sigma) = x \mid \mathrm{DNF}(I^\sigma) = m\right].$$

Using this fact and the assumption that we consider a perfect-packing instance it is possible to estimate the random-order ratio in terms of the probability that the overshoot is bounded from above by $1/2$:

▶ **Lemma 70.** *Let $\mathcal{I}$ be a set of items satisfying the assumptions, and let* $\mathcal{E} = \{I^\sigma \mid R_1(I^\sigma) \leq 1/2\}$*. Then it holds*

$$\frac{\mathbb{E}\left[\mathrm{DNF}(I^\sigma)\right]}{\mathrm{OPT}(\mathcal{I})} \geq \frac{1}{2} + \frac{1}{4} \cdot \mathbb{P}\left[\mathcal{E}\right] - \frac{1}{\mathrm{OPT}(\mathcal{I})}. \tag{19}$$

Now, it remains to show a lower bound for the probability of $\mathcal{E}$, which is independent of the considered set $\mathcal{I}$.

▶ **Lemma 71.** *Let ϵ greater than zero be arbitrary, and let $\mathcal{I}$ be an arbitrary set of items satisfying the assumptions with* $\mathrm{OPT}(\mathcal{I})$ *large enough. Then, we have*

$$\mathbb{P}\left[\mathcal{E}\right] \geq \frac{7}{10}e^{-3} \cdot \left(1 - \exp\left(-\frac{121}{420}\right)\right) - \epsilon.$$

Plugging this into (19) we obtain the lower bound given in Theorem 67, that is the estimate

$$\mathrm{RR}(\mathrm{DNF}) \geq \frac{1}{2} + \frac{7}{40}e^{-3}\left(1 - \exp\left(-\frac{121}{420}\right)\right) \geq 0.502.$$

This is complemented by an upper bound, which improves the result given in [17]:

▶ **Lemma 72.** *Let $\mathcal{D}$ denote the set of all distributions on* $[0,1]$. *Then, we have*

$$\mathrm{APR}(\mathrm{DNF}, \mathcal{D}) \leq 2/3.$$

Using Lemma 32 this transfers to an upper bound for the random-order ratio. This completes the proof of Theorem 67.

On the other hand, the algorithm Dual-Harmonic does not benefit from random input.

▶ **Proposition 73.** *Let $M \in \mathbb{N}$ with $M \geq 2$. Then it holds* $\mathrm{APR}(\mathrm{DH}_M, \mathcal{D}) = 1/2$.

4.2.2 Deferred Proofs

4.2.2.1 Proof of Lemma 68

▶ **Lemma 68.** *Let I be an instance and let ℓ denote the number of large items in I. If $\ell < 1/4 \cdot \mathrm{OPT}(I)$, then we have* $\mathrm{DNF}(I) \geq \frac{7}{12}\,\mathrm{OPT}(I) - 2$.

Proof of Lemma 68. The proof is based on the observation that the overshoot of bins that contain only small items is bounded from above by $1/2$. As the number of large items is smaller than $1/4\,\mathrm{OPT}(\mathcal{I})$ there can be at most $1/4\,\mathrm{OPT}(\mathcal{I})$ many bins with overshoot 1. Hence, there are small items of total mass at least $S(\mathcal{I}) - 1/4\,\mathrm{OPT}(\mathcal{I}) \cdot 2 = 1/2 \cdot S(\mathcal{I})$ remaining. These items cover at least $1/2 \cdot S(\mathcal{I})/(3/2) = 1/3\,\mathrm{OPT}(\mathcal{I})$ many bins.

Hence, the total number of covered bins is at least

$$\left\lfloor \frac{1}{4}\,\mathrm{OPT}(\mathcal{I}) \right\rfloor + \left\lfloor \frac{1}{3}\,\mathrm{OPT}(\mathcal{I}) \right\rfloor \geq \frac{7}{12}\,\mathrm{OPT}(\mathcal{I}) - 2.$$

◀

4.2.2.2 Proof of Lemma 69

▶ **Lemma 69.** *Let $x \in [0,1]$ and $m \in \{1, \ldots, \mathrm{OPT}(\mathcal{I})\}$ with $\mathbb{P}\left[\mathrm{DNF}(I^\sigma) = m\right] > 0$. Then, for $i, j \leq m$ we have*

$$\mathbb{P}\left[R_i(I^\sigma) = x \mid \mathrm{DNF}(I^\sigma) = m\right] = \mathbb{P}\left[R_j(I^\sigma) = x \mid \mathrm{DNF}(I^\sigma) = m\right].$$

Proof of Lemma 69. Without loss of generality we assume that all item sizes are different, otherwise we distinguish them by their index. Let $\mathcal{I} = \{a_1, \ldots, a_n\}$ denote the set of items. Moreover, let $I^\sigma = (A_1, \ldots, A_n)$ with $A_i = a_{\sigma(i)}$ be a random instance.

Let $m \in \{1, \ldots, \mathrm{OPT}(\mathcal{I})\}$ with $\mathbb{P}[\mathrm{DNF}(I^\sigma) = m] > 0$. Furthermore, let $I = (v_1, \ldots, v_n)$ denote a vector with $v_i \in \mathcal{I}$ for all $i \in [n]$ and $\mathbb{P}[I^\sigma = I \mid \mathrm{DNF}(I^\sigma) = m] > 0$. We notice that there is only one possible realization of I^σ such that $I^\sigma = I$. Therefore, it follows that $\mathbb{P}[I^\sigma = I \mid \mathrm{DNF}(I^\sigma) = m] = \frac{1}{\#\{\sigma : \mathrm{DNF}(I^\sigma) = m\}}$.

Let $u \in [m]$ be arbitrary. Since $\mathrm{DNF}(I) = m$ there exist indices i, j and k such that $v_1, \ldots, v_i$ are the items that are used to cover the first bin, and $v_j, \ldots, v_k$ the items that are used to cover the u-th bin. Now let

$$I' = (v_j, \ldots, v_k, v_{i+1}, \ldots, v_{j-1}, v_1, \ldots, v_i, v_{k+1}, \ldots, v_n).$$

I' is the instance we obtain from I by changing the configuration of the first and the u-th bin. Then, $\mathrm{DNF}(I')$ is also equal to m and we have $\mathbb{P}[\mathrm{DNF}(I^\sigma) = I \mid \mathrm{DNF}(I^\sigma) = m] = \mathbb{P}[\mathrm{DNF}(I^\sigma) = I' \mid \mathrm{DNF}(I^\sigma) = m]$.

So, we have shown that the bin levels of the first and the u-th bin, where u is arbitrary, are conditionally identically distributed. It follows immediately that the same is true for the overshoot. ◀

4.2.2.3 Proof of Lemma 70

▶ **Lemma 70.** *Let $\mathcal{I}$ be a set of items satisfying the assumptions, and let $\mathcal{E} = \{I^\sigma \mid R_1(I^\sigma) \leq 1/2\}$. Then it holds*

$$\frac{\mathbb{E}[\mathrm{DNF}(I^\sigma)]}{\mathrm{OPT}(\mathcal{I})} \geq \frac{1}{2} + \frac{1}{4} \cdot \mathbb{P}[\mathcal{E}] - \frac{1}{\mathrm{OPT}(\mathcal{I})}. \tag{19}$$

Proof of Lemma 70. Because of Assumption 6 the following identity is true for every realization of I^σ:

$$\mathrm{OPT}(\mathcal{I}) = S(I^\sigma) = \mathrm{DNF}(I^\sigma) + \sum_{i=1}^{\mathrm{DNF}(I^\sigma)} R_i(I^\sigma) + W(I^\sigma).$$

Applying expected values to both sides and estimating $W(I^\sigma)$ by 1 it follows that

$$\mathrm{OPT}(\mathcal{I}) \leq \mathbb{E}[\mathrm{DNF}(I^\sigma)] + \mathbb{E}\left[\sum_{i=1}^{\mathrm{DNF}(I^\sigma)} R_i(I^\sigma)\right] + 1. \tag{20}$$

Let $Q := \{j \in \{1, \ldots, \mathrm{OPT}(\mathcal{I})\} : \mathbb{P}[\mathrm{DNF}(I^\sigma) = j] > 0\}$. Now, we use the law of total expectation and apply Lemma 69. It follows that

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{\mathrm{DNF}(I^\sigma)} R_i(I^\sigma)\right] &= \sum_{j \in Q} \mathbb{E}\left[\sum_{i=1}^{\mathrm{DNF}(I^\sigma)} R_i(I^\sigma) \,\middle|\, \mathrm{DNF}(I^\sigma) = j\right] \cdot \mathbb{P}[\mathrm{DNF}(I^\sigma) = j] \\
&= \sum_{j \in Q} \mathbb{E}\left[\sum_{i=1}^{j} R_i(I^\sigma) \,\middle|\, \mathrm{DNF}(I^\sigma) = j\right] \cdot \mathbb{P}[\mathrm{DNF}(I^\sigma) = j] \\
&= \sum_{j \in Q} \sum_{i=1}^{j} \mathbb{E}[R_i(I^\sigma) | \mathrm{DNF}(I^\sigma) = j] \cdot \mathbb{P}[\mathrm{DNF}(I^\sigma) = j]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j \in Q} \sum_{i=1}^{j} \mathbb{E}\left[R_1(I^\sigma) \middle| \operatorname{DNF}(I^\sigma) = j\right] \cdot \mathbb{P}\left[\operatorname{DNF}(I^\sigma) = j\right] \\
&= \sum_{j \in Q} \mathbb{E}\left[j \cdot R_1(I^\sigma) \middle| \operatorname{DNF}(I^\sigma) = j\right] \cdot \mathbb{P}\left[\operatorname{DNF}(I^\sigma) = j\right] \\
&= \sum_{j \in Q} \mathbb{E}\left[\operatorname{DNF}(I^\sigma) \cdot R_1(I^\sigma) \middle| \operatorname{DNF}(I^\sigma) = j\right] \cdot \mathbb{P}\left[\operatorname{DNF}(I^\sigma) = j\right] \\
&= \mathbb{E}\left[\operatorname{DNF}(I^\sigma) \cdot R_1(I^\sigma)\right].
\end{aligned}
$$

Combining this with (20), we obtain

$$
\begin{aligned}
\operatorname{OPT}(\mathcal{I}) &\leq \mathbb{E}\left[\operatorname{DNF}(I^\sigma)\right] + \mathbb{E}\left[\operatorname{DNF}(I^\sigma) \cdot R_1(I^\sigma)\right] + 1 \\
&= \mathbb{E}\left[\operatorname{DNF}(I^\sigma) \cdot (1 + R_1(I^\sigma)) \middle| R_1(I^\sigma) \leq \frac{1}{2}\right] \cdot \mathbb{P}\left[R_1(I^\sigma) \leq \frac{1}{2}\right] \\
&\quad + \mathbb{E}\left[\operatorname{DNF}(I^\sigma) \cdot (1 + R_1(I^\sigma)) \middle| R_1(I^\sigma) > \frac{1}{2}\right] \cdot \mathbb{P}\left[R_1(I^\sigma) > \frac{1}{2}\right] + 1 \\
&\leq \frac{3}{2}\mathbb{E}\left[\operatorname{DNF}(I^\sigma) \middle| R_1(I^\sigma) \leq \frac{1}{2}\right] \cdot \mathbb{P}\left[R_1(I^\sigma) \leq \frac{1}{2}\right] \\
&\quad + 2\mathbb{E}\left[\operatorname{DNF}(I^\sigma) \middle| R_1(I^\sigma) > \frac{1}{2}\right] \cdot \mathbb{P}\left[R_1(I^\sigma) > \frac{1}{2}\right] + 1 \\
&= 2\mathbb{E}\left[\operatorname{DNF}(I^\sigma)\right] - \frac{1}{2}\mathbb{E}\left[\operatorname{DNF}(I^\sigma) \middle| R_1(I^\sigma) \leq \frac{1}{2}\right] \cdot \mathbb{P}\left[R_1(I^\sigma) \leq \frac{1}{2}\right] + 1 \\
&\leq 2\mathbb{E}\left[\operatorname{DNF}(I^\sigma)\right] - \frac{1}{2}\operatorname{OPT}(\mathcal{I}) \cdot \mathbb{P}\left[R_1(I^\sigma) \leq \frac{1}{2}\right] + 1.
\end{aligned}
$$

Therefore, it follows that

$$
\frac{\mathbb{E}\left[\operatorname{DNF}(I^\sigma)\right]}{\operatorname{OPT}(\mathcal{I})} \geq \frac{1}{2} + \frac{1}{4}\mathbb{P}\left[R_1(I^\sigma) \leq \frac{1}{2}\right] - \frac{1}{\operatorname{OPT}(\mathcal{I})}.
$$

◀

4.2.2.4 Proof of Lemma 71

▶ **Lemma 71.** *Let ϵ greater than zero be arbitrary, and let $\mathcal{I}$ be an arbitrary set of items satisfying the assumptions with* $\operatorname{OPT}(\mathcal{I})$ *large enough. Then, we have*

$$
\mathbb{P}\left[\mathcal{E}\right] \geq \frac{7}{10}e^{-3} \cdot \left(1 - \exp\left(-\frac{121}{420}\right)\right) - \epsilon.
$$

Proof of Lemma 71. Let ℓ denote the number of large, and m the number of small items in $\mathcal{I}$. We set $r := m/\ell$, and we can assume without loss of generality that r is a positive natural number, otherwise we add items of size zero to $\mathcal{I}$. Those added items do not effect the number of covered bins.

Let $g_1 \geq g_2 \geq \ldots \geq g_\ell$ denote the sizes of the large items in $\mathcal{I}$, and we set $s^\star := 1 - g_{\lceil 3\ell/10 \rceil}$. We are now interested in the following three events

$\mathcal{F}_1 := \{$the size of the first drawn large item in I^σ is at least $1 - s^\star\}$

$\mathcal{F}_2 := \{$there is at most one large item among the first $3r + 1$ drawn items in $I^\sigma\}$

$\mathcal{F}_3 := \{$the sum of the first $3r$ small items in I^σ is at least $s^\star\}$.

Let $\mathcal{F} := \mathcal{F}_1 \cap \mathcal{F}_2 \cap \mathcal{F}_3$. We set $\mathcal{F}_a := \mathcal{F} \cap \{I^\sigma \mid R_1(I^\sigma) \leq 1/2\}$ and $\mathcal{F}_b := \mathcal{F} \cap \{I^\sigma \mid R_1(I^\sigma) > 1/2\}$. It holds true that $\mathbb{P}[\mathcal{F}_a] \geq 2\mathbb{P}[\mathcal{F}_b]$ as the following argument shows: If $\mathcal{F}_b$ occurs we cover the bin with a large item, which arrives at the end. Before, there are at least two small items, otherwise the sum cannot be greater than $3/2$. Now we look at the permutations, where the small items arrive in the same order, but the large item appears at the first or second position. In both cases, we cover the bin with overshoot at most $1/2$. Furthermore, all permutations have the same probability. So the statement follows.

Hence, the following holds:

$$\mathbb{P}[\mathcal{E}] \geq \mathbb{P}[\mathcal{F}_a] \geq \frac{2}{3}\mathbb{P}[\mathcal{F}].$$

So, in the remaining part of the proof we will establish a lower bound for $\mathbb{P}[\mathcal{F}]$ that is independent of $\mathcal{I}$. We observe that the three events $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{F}_3$ are independent, therefore it follows that $\mathbb{P}[\mathcal{F}] = \mathbb{P}[\mathcal{F}_1] \cdot \mathbb{P}[\mathcal{F}_2] \cdot \mathbb{P}[\mathcal{F}_3]$.

Due to the definition of $s^\star$ it is clear that $\mathbb{P}[\mathcal{F}_1] \geq 3/10$. So, we continue with considering $\mathbb{P}[\mathcal{F}_2]$: We distinguish two cases: In the first case, all $3r+1$ items drawn are small. In the second case, there is exactly one large item. We start estimating the first case. Let $\epsilon' > 0$ arbitrary. It follows from Assumption 7 that if $\mathrm{OPT}(\mathcal{I})$ is large enough, we have $3/\ell \leq \epsilon'$. So we obtain

$$\prod_{i=0}^{3r}\left(1-\frac{\ell}{m+\ell-i}\right) \geq \left(1-\frac{\ell}{m+\ell-3r}\right)^{3r+1} = \left(1-\frac{1}{r+1-3r/\ell}\right)^{3r+1}$$
$$\geq \left(1-\frac{1}{1+r\cdot(1-\epsilon')}\right)^{3r}\cdot\left(1-\frac{1}{1+r\cdot(1-\epsilon')}\right) \geq \exp\left(-\frac{3}{1-\epsilon'}\right)\cdot\frac{1-\epsilon'}{2-\epsilon'}.$$

Now assume that there is exactly one large item. Then, it follows

$$(3r+1)\cdot\left(\prod_{i=0}^{3r-1}\left(1-\frac{\ell}{m+\ell-i}\right)\right)\cdot\frac{\ell}{m-3r}$$
$$\geq \ell\cdot\frac{3r+1}{m-3r}\cdot\exp\left(-\frac{3}{1-\epsilon'}\right) \geq \frac{3}{1-\epsilon'}\cdot\exp\left(-\frac{3}{1-\epsilon'}\right).$$

Summing up both probabilities, we see that for every $\epsilon'' > 0$ we have

$$\mathbb{P}[\mathcal{F}_2] \geq \frac{7}{2}e^{-3} - \epsilon'',$$

if $\mathrm{OPT}(\mathcal{I})$ is large enough.

It remains to give a bound for $\mathbb{P}[\mathcal{F}_3]$: We set the sizes of all small items that are packed in an optimal solution in bins that contain only small items or one of the items with size $g_1, \dots, g_{\lceil 3\ell/10\rceil -1}$ to zero. Furthermore, we assume that the total size of small items in each of the bins containing one of the large items $g_{\lceil 3\ell/10\rceil}, \dots, g_\ell$ is equal to $s^\star$. This only decreases the probability we want to bound from below. Let X_i be the size of the i-th drawn small item. We can think of X_i as a sample according to sampling without replacement from the multiset of small items. Then, we have $\mathbb{E}[X_i] \geq \frac{7}{10}\ell s^\star/m$ and $\mathbb{E}\left[X_i^2\right] \leq \left(\frac{7}{10}\ell+1\right)\cdot(s^\star)^2/m$. Let X denote the sum of the first $3r$ small items. Then, we have

$$\mathbb{E}[X] \geq 3\frac{m}{\ell}\cdot\frac{7}{10}\ell s^\star/m \geq \frac{21}{10}s^\star.$$

Set $t = \frac{11}{10}s^\star$. Now, applying Bernstein's inequality we obtain

$$\mathbb{P}[X \leq s^\star] \leq \mathbb{P}\left[X \leq \mathbb{E}[X] - \frac{11}{10}s^\star\right] \leq \exp\left(-\frac{(11s^\star/10)^2}{2\cdot 3r \cdot \left[\frac{7}{10}\cdot\frac{\ell}{m}(s^\star)^2 + \frac{(s^\star)^2}{m}\right]}\right)$$
$$= \exp\left(-\frac{\frac{121}{100}}{\frac{42}{10}+\frac{6}{\ell}}\right) \xrightarrow{\ell\to\infty} \exp\left(-\frac{121}{420}\right).$$

Combining the results for $\mathbb{P}[\mathcal{F}_1]$, $\mathbb{P}[\mathcal{F}_2]$ and $\mathbb{P}[\mathcal{F}_3]$ we see that for every $\epsilon > 0$ we have

$$\mathbb{P}[\mathcal{E}] \geq \frac{2}{3}\cdot\mathbb{P}[\mathcal{F}_1]\cdot\mathbb{P}[\mathcal{F}_2]\cdot\mathbb{P}[\mathcal{F}_3] \geq \frac{7}{10}e^{-3}\cdot\left(1-\exp\left(-\frac{121}{420}\right)\right) - \epsilon,$$

if $\mathrm{OPT}(\mathcal{I})$ is large enough. ◀

4.2.2.5 Proof of Lemma 72

▶ **Lemma 72.** *Let $\mathcal{D}$ denote the set of all distributions on $[0,1]$. Then, we have*

$$\mathrm{APR}(\mathrm{DNF},\mathcal{D}) \leq 2/3.$$

Proof Lemma 72. We will show that for every $\epsilon > 0$ we can find a distribution F such that $\mathrm{APR}(\mathrm{DNF},F) \leq 3/2 + \epsilon$. Let $F_{m,k}$ denote the uniform distribution on the set of items

$$\mathcal{I} = \left\{\frac{1}{k}, 1-\frac{1}{k}, \left(\frac{1}{k}\right)^2, 1-\left(\frac{1}{k}\right)^2, \ldots, \left(\frac{1}{k}\right)^m, 1-\left(\frac{1}{k}\right)^m\right\}.$$

Since we are dealing with a maximization problem we can apply Lemma 31. So, in the following we will separately bound $\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right]$ from above, and $\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]$ from below.

To analyze $\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right]$ we will use Markov chains. We can describe the behavior of $\mathrm{DNF}(I_n^F)$ by considering the Markov chain on the set of possible bin levels of $\mathrm{DNF}(I_n^F)$. Here, we subsume bin level zero and all bin levels greater than or equal to one to a special state $(*)$.

This Markov chain is irreducible due to its construction and therefore there exists a stationary measure π. Let $\mathrm{DNF}(u,I)$ denote the number of covered bins, if Dual-Next-Fit starts with bin level u. Furthermore, we set $\mathrm{DNF}((*),I) := \mathrm{DNF}(0,I) = \mathrm{DNF}(I)$.

As Dual-Next-Fit is a monotone algorithm we have $\mathrm{DNF}(I) \leq \mathrm{DNF}(u,I)$ for an arbitrary bin level u. Moreover, if the starting bin level L is distributed according to π, then we have

$$\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right] \leq \mathbb{E}_{L\sim\pi}\left[\mathrm{DNF}(L,I_n^F)\right] = n\cdot\pi((*)) = \frac{n}{\mathbb{E}[T_F]}.$$

Here, T_F denotes the number of items we need to cover a bin, starting from state $(*)$, that is with an empty bin. This follows from the property that π is a stationary distribution, which is explained in Proposition 21 in Section 2.2.

Now, we show that for every $\epsilon' > 0$ there exist parameters m and k such that $\mathbb{E}\left[T_{F_{m,k}}\right] \geq 3-\epsilon'$. It holds

$$\mathbb{E}\left[T_{F_{m,k}}\right] = \sum_{i=0}^{\infty}\mathbb{P}\left[T_{F_{m,k}} > i\right] \geq 2 + \sum_{i=2}^{k-1}\mathbb{P}\left[T_{F_{m,k}} > i\right].$$

Simple counting yields for $i \geq 2$

$$\mathbb{P}\left[T_{F_{m,k}} > i\right] = \frac{m^i}{(2m)^i} + \frac{1}{(2m)^i}\sum_{j=2}^{m} i\cdot(j-1)^{i-1} = \frac{1}{2^i} + \frac{i}{2^i m^i}\sum_{j=1}^{m-1} j^{i-1}$$

$$\geq \frac{1}{2^i} + \frac{i}{2^i m^i} \cdot \int_0^{m-1} x^{i-1}\, dx = \frac{1}{2^i} \cdot \left[1 + \left(1 - \frac{1}{m}\right)^i\right].$$

Therefore, if we choose at first k, and then m large enough, we obtain

$$\mathbb{E}\left[T_{F_{m,k}}\right] \geq 2 + \sum_{i=2}^{k-1} \frac{1}{2^i} \cdot \left[1 + \left(1 - \frac{1}{m}\right)^i\right] \geq 3 - \epsilon'.$$

Hence, for every $\epsilon' > 0$ we can find a distribution $F_{m,k}$ with

$$\mathbb{E}\left[\mathrm{DNF}(I_n^{F_{m,k}})\right] \leq \mathbb{E}_{L\sim\pi}\left[\mathrm{DNF}(L, I_n^{F_{m,k}})\right] \leq \frac{n}{3 - \epsilon'}. \tag{21}$$

Now, we give a lower bound for $\mathbb{E}\left[\mathrm{OPT}(I_n^{F_{m,k}})\right]$: Applying Hoeffding's inequality for each item size and applying a union bound we obtain that with probability at least $1 - 2m\exp\left(-2\log(n)^{3/2}\right)$ there are at least $\frac{n}{2m} - \sqrt{n}\log(n)^{3/4}$ many items of each type. Then, $\mathrm{OPT}(I_n^{F_{m,k}})$ is lower bounded by $n/2 - 2m\log(n)^{3/4}$. Therefore, it follows that

$$\mathbb{E}\left[\mathrm{OPT}(I_n^{F_{m,k}})\right] \geq \left(1 - 2m\exp\left(-2\log(n)^{3/2}\right)\right) \cdot \left(n/2 - 2m\log(n)^{3/4}\right). \tag{22}$$

Combining (21) and (22) it follows that if n tends to infinity we can find for every $\epsilon > 0$ a distribution $F_{m,k}$ such that

$$\limsup_{n\to\infty} \frac{\mathbb{E}\left[\mathrm{DNF}(I_n^{F_{m,k}})\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^{F_{m,k}})\right]} \leq \frac{2}{3} + \epsilon.$$

◀

4.2.2.6 Proof of Proposition 73

▶ **Proposition 73.** *Let $M \in \mathbb{N}$ with $M \geq 2$. Then it holds* $\mathrm{APR}(\mathrm{DH}_M, \mathcal{D}) = 1/2$.

Proof of Proposition 73. Since bin covering is a maximization problem, we have $\frac{\mathrm{DH}_M(I)}{\mathrm{OPT}(I)} \leq 1$. Hence, we can apply Lemma 31. Let $m \in \mathbb{N}$. Let F be given by the uniform distribution on $\mathcal{I} = \{1/m, 1 - 1/m\}$. We call items of size $1/m$ small and of size $1 - 1/m$ big. Let $sm(I_n^F)$ denote the number of small items and $b(I_n^F)$ the number of big items in a random instance. Then we have

$$\mathbb{E}\left[\mathrm{DH}_M(I_n^F)\right] \leq \mathbb{E}\left[b(I_n^F)\right]/2 + \mathbb{E}\left[sm(I_n^F)\right]/m = n/4 + n/(2m).$$

Furthermore, we can pair the small and big items. Therefore, we have $\mathrm{OPT}(I_n^F) \geq \min\{sm(I_n^F), b(I_n^F)\}$. Using Hoeffding's inequality two times and applying a union bound we see that

$$\mathbb{P}\left[\min\{sm(I_n^F), b(I_n^F)\} \geq (n/2 - \sqrt{n\log(n)})\right] \geq 1 - \frac{4}{n^2}.$$

Therefore

$$\frac{\mathbb{E}\left[\mathrm{DH}_M(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \leq \frac{n/4 + n/(2m)}{(n/2 - \sqrt{n\log(n)}) \cdot (1 - 4/n^2)}.$$

Since we can choose m arbitrary large, the statement follows for the average performance ratio. ◀

5 Analysis of Selected Heuristics for Class-constrained Bin Packing and Bin Covering

	lower bound	upper bound
$\mathrm{CR}(A)$	2 [7]	2.635 [38]
$\mathrm{RR}(A)$	$1.\bar{1}$	?
$\mathrm{APR}(A, \mathcal{D})$	1	1

Figure 17 Summary of upper and lower bounds for class-constrained bin packing, including the results obtained in Section 3.

	l. bound	u. bound
$\mathrm{CR}(\mathrm{CS}\,[A])$	2 [77]	3 [38]
$\mathrm{RR}(\mathrm{CS}\,[A])$	?	?
$\mathrm{APR}(\mathrm{CS}\,[A], \mathcal{D})$	?	?

	l. bound	u. bound
$\mathrm{CR}(\mathrm{CS}\,[A])$	2 [77]	3 [38]
$\mathrm{RR}(\mathrm{CS}\,[A])$	**2**	?
$\mathrm{APR}(\mathrm{CS}\,[A], \mathcal{D})$	$\boldsymbol{\approx 1.691}$	$\boldsymbol{\approx 1.691 + \epsilon}$

Figure 18 Comparison between previously given bounds for algorithms based on the ColorSets-approach for class-constrained bin packing with general item sizes on the left hand side and our results on the right hand side.

	l. bound	u. bound
$\mathrm{CR}(A)$	$\Omega\left(\frac{1}{k}\right)$ [37]	$\mathcal{O}\left(\frac{1}{\log(k)}\right)$ [37]
CR(DNF)	0 [37]	0 [37]
RR(DNF)	?	?
$\mathrm{APR}(\mathrm{DNF}, \mathcal{D})$	?	?
CR(FF2)	$\frac{1}{B}$ [37]	$\frac{1}{B}$ [37]
RR(FF2)	?	?
$\mathrm{APR}(\mathrm{FF2}, \mathcal{D})$	?	?

	l. bound	u. bound
$\mathrm{CR}(A)$	$\Omega\left(\frac{1}{k}\right)$ [37]	$\mathcal{O}\left(\frac{1}{\log(k)}\right)$ [37]
CR(DNF)	0 [37]	0 [37]
RR(DNF)	$\boldsymbol{\Omega\left(\frac{1}{\log(k)}\right)}$	$\boldsymbol{\mathcal{O}\left(\frac{1}{\log(k)}\right)}$
$\mathrm{APR}(\mathrm{DNF}, \mathcal{D})$	$\boldsymbol{\Omega\left(\frac{1}{\log(k)}\right)}$	$\boldsymbol{\mathcal{O}\left(\frac{1}{\log(k)}\right)}$
CR(FF2)	$\frac{1}{B}$ [37]	$\frac{1}{B}$ [37]
RR(FF2)	**1**	**1**
$\mathrm{APR}(\mathrm{FF2}, \mathcal{D})$	**1**	**1**

Figure 19 Comparison between previously given bounds for online class-constrained bin covering with unit sized items on the left hand side and our results on the right hand side.

In this part we consider selected heuristics for class-constrained bin packing and bin covering. We have already investigated the complexity of class-constrained bin packing in Section 3. There, we found out that the complexity of this problem differs according to the two studied probabilistic performance measures. While it is possible to construct an optimal algorithm with respect to the average performance ratio, we obtained a lower bound of $10/9$ for the random-order ratio.

For class-constrained bin packing we consider the case of general item sizes, that is the item sizes are arbitrary in $[0, 1]$, and the case of unit sized items, that is all items have size 1 and the bin capacity is given by an integer B. We will find out that the phenomenon of the different complexity also plays an important role for the heuristics under consideration.

When we consider the dual problem, that is class-constrained bin covering we do not observe different behaviors. Here, the analyzed heuristics benefit notably from both randomized settings. A comparison between some of our results and corresponding results in literature is given in Figure 17, 18 and 19.

Since some of the proofs are technical, we will present the statements (and sometimes ideas of the proof) in one section, and postpone the proofs to a following section.

5.1 Results

We start with studying class-constrained bin packing with general item sizes. We remember that in the worst case there exists a lower bound of 2 with respect to the competitive ratio and that the best known online algorithm in this setting has a competitive ratio of approximately 2.635 [38].

The proofs in this section are mostly based on two techniques:

- The first important observation is that if the items are independent and identically distributed, then the expected number of different drawn colors can grow only sublinearly in terms of the number of drawn items.
- The second technique is the following: Often our designed instances contain many small items. They are chosen in such a way that their total mass is negligible in comparison to the value of the optimal solution, but they determine a certain order of the first arrival of the colors with high probability.

A possible way to reuse online algorithms for classical bin packing is the ColorSets-approach: Here, we partition the items into groups of size k according to the order of their first arrival. Then, all items with colors that belong to the same group, are packed by an online algorithm A that is designed for classical bin packing. In the following we will analyze this approach with respect to the average performance ratio and we will show that there is a connection to the lower bound for bounded-space online algorithms for classical bin packing $h_\infty \approx 1.691$ given by Lee and Lee [58].

▶ **Theorem 74.** *Let $\mathcal{D}$ denote the set of all distributions on $[0,1] \times \mathbb{N}$, and let $\epsilon > 0$ be arbitrary. Choosing M sufficiently large, it holds*

$$\mathrm{APR}(\mathrm{CS}\left[\mathrm{HARMONIC}_M\right], \mathcal{D}) \leq h_\infty + \epsilon.$$

Furthermore, let $\epsilon > 0$ be arbitrary. Then, there exists a parameter k for class-constrained bin packing such that for every online algorithm A, which is intended for classical bin packing, it holds

$$\mathrm{APR}(\mathrm{CS}\left[A\right], \mathcal{D}) \geq h_\infty - \epsilon.$$

So we see that if the items are sampled independently and identically distributed it is possible to beat the lower bound of two for the competitive ratio. The proof of the upper bound is based on the sublinear growth of different colors. For the proof of the lower bound we introduce many small items that determine the order of the first arrival of colors with high probability. Then, it follows that (apart from the small items) in the first group items of size $1/2 + \epsilon$, in the second group items of size $1/3 + \epsilon$, in the third group items of size $1/7 + \epsilon$, ... are packed. Hence, we encounter a situation as in the proof of the lower bound for bounded-space online algorithms for classical bin packing given in [58].

In the random-order model things are more complicated: We can show that it is not possible for ColorSets-based algorithms and First-Fit to obtain a performance of h_∞.

▶ **Proposition 75.** *Let A be an arbitrary online algorithm for classical bin packing. If $k = 2$, it holds $\mathrm{RR}(\mathrm{CS}\left[A\right]) \geq 2$. Moreover, for every ϵ greater than zero, there exists a parameter k such that $\mathrm{RR}(\mathrm{FF}) \geq 2 - \epsilon$.*

This proof is based again on the fact that we can determine the order of the first arrival of the colors with high probability by introducing small items, whose sizes are negligible for the optimal solution. The consequence is that algorithms based on the ColorSets-approach must pack in almost all groups items of total size slightly larger than 1. Here, we also exploit the fact that the adversary in the random-order model is more powerful than in the case that the items are drawn independently and identically distributed: In the random-order model it is possible to restrict the total size of items of one color, while this quantity is always growing if the items are sampled independently and identically distributed.

Now we consider the case of unit sized items. We observe the same behavior of algorithms as in the case of general item sizes. If the items are drawn independently and identically distributed a large class of natural algorithms performs asymptotically optimal, but in the random-order model their performance is worse.

▶ **Proposition 76.** *We consider class-constrained bin packing with unit sized items with parameters B and k. Let $\mathcal{D}$ denote the set of all possible distributions, and let A be* CS [NF] *or any online algorithm that opens a new bin only if it is forced, then it holds*

$$\mathrm{APR}(A, \mathcal{D}) = 1.$$

This proof is again based on the fact that the expected number of different drawn colors grows sublinearly in the number of items. The performance according to the random-order ratio is again weaker:

▶ **Proposition 77.** *For every ϵ greater than zero there exist parameters B and k for class-constrained bin packing with unit sized items such that* $\mathrm{RR}(\mathrm{CS}\,[\mathrm{NF}]) \geq 2 - \epsilon$.

▶ **Proposition 78.** *For every ϵ greater than zero there exist parameters B and k for class-constrained bin packing with unit sized items such that* $\mathrm{RR}(\mathrm{FF}) \geq 1.5 - \epsilon$.

Finally, we point out that bounded-space algorithms perform poorly for class-constrained bin packing, even on random input with unit sized items. This is in contrast to classical bin packing.

▶ **Proposition 79.** *Consider class-constrained bin packing with unit sized items and parameters B and k. Let A be an arbitrary bounded-space online algorithm. Then we have* $\mathrm{RR}(A), \mathrm{APR}(A, \mathcal{D}) \in \Omega(B/k)$.

Now, we turn to class-constrained bin covering. Up to now this problem was only studied in the setting of unit sized items. It was shown that no online algorithm can obtain a competitive ratio better than of order $\log(k)^{-1}$, and there exists a heuristic with competitive ratio of order $1/k$. Moreover, the heuristic FF2, which is based on the First-Fit-approach, was shown to be $1/B$-competitive. All these results can be found in [37].

We start with a slightly improved upper bound for the competitive ratio of online algorithms for class-constrained bin covering with unit sized items. This proof uses the same technique as in [37], but adjusts the choice of scenarios.

▶ **Proposition 80.** *The competitive ratio of any deterministic online algorithm for class-constrained bin covering with unit sized items is at most $\left(H_{k-1} + 1 - \frac{k-1}{B}\right)^{-1}$. If $B = k$ this yields an upper bound of H_k^{-1}. The same is true for randomized algorithms.*

Furthermore, we show that bounded-space algorithms behave poorly according to the competitive ratio. This is the same behavior as in class-constrained bin packing, but in contrast to classical bin covering.

▶ **Proposition 81.** *Let A be a bounded-space algorithm for class-constrained bin covering. Then, it holds* $\mathrm{CR}(A) = 0$.

Now, we analyze the performance of the heuristics Dual-Next-Fit and FF2 with respect to both probabilistic performance measures. Before starting we have to introduce a new class of distributions: We remember that a distribution F was introduced as a pair of a multiset of items $\mathcal{I}$ and a probability mass function/density function p. In case that $\mathcal{I}$ is finite and p is the uniform distribution on $\mathcal{I}$ we call this distribution a *discrete distribution*. Furthermore, if $S(\mathcal{I}) = \mathrm{OPT}(\mathcal{I})$ we call the distribution a *discrete perfect-packing distribution*.

We start with the simple 1-bounded space algorithm Dual-Next-Fit.

▶ **Theorem 82.** *For class-constrained bin covering with unit sized items and parameters k and B it holds* $\mathrm{RR}(\mathrm{DNF}) \in \Theta(\log(k)^{-1})$. *In the case of general item sizes let $\mathcal{D}$ denote the set of all discrete perfect-packing distributions. Then we have* $\mathrm{APR}(\mathrm{DNF}, \mathcal{D}) \in \Theta(\log(k)^{-1})$.

▶ Remark. Dual-Next-Fit is a *monotone* algorithm. That is, decreasing the size of items or removing them from the input, only decreases the number of covered bins. Because of the monotonicity it is possible to show that the result concerning the average performance ratio carries over to the case, where $\mathcal{D}$ is the set of all discrete distributions.

Surprisingly it turns out that the heuristic FF2, which is based on the First-Fit-approach, is even optimal if unit sized items arrive in random order:

▶ **Theorem 83.** *For class-constrained bin covering with unit sized items and parameters k and B it holds* $\mathrm{RR}(\mathrm{FF2}) = 1$.

So we see that in class-constrained bin covering with unit sized items the heuristics under consideration benefit strongly from randomized input: While the 1-bounded-space heuristic Dual-Next-Fit does not satisfy any guarantees with respect to the competitive ratio, its random-order ratio is of the same order as the upper bound for algorithms in the worst case. The algorithm FF2, which is based on a First-Fit-approach, is even optimal if the items arrive in random order. This is also in contrast to class-constrained bin packing: In Proposition 78 we have seen that a First-Fit approach for the dual version of the problem cannot be optimal if the items arrive in random order.

The idea of the proof of Theorem 82 and 83 is as follows: At first we exploit the fact that both algorithms are *monotone*. This allows us to deal only with perfect-packing instances, that is, it holds $\mathrm{OPT}(I) = S(I)$. Afterwards, we subdivide the input into smaller subinstances for which we can assume that the items are drawn independently and identically distributed from a distribution F. Then, we apply the tools from the field of Markov chains.

For Dual-Next-Fit, similar to the analysis of Dual-Next-Fit for classical bin covering, we estimate the number of items we need to cover a bin. In order to do this we exploit a connection to the *Coupon-collectors problem*. Then, we use the connection between the number of items we need to cover a bin and the long-time average number of opened bins.

For FF2 we construct a comparison chain, which simulates the behavior of the algorithm on an instance which is worse for the algorithm. Afterwards, we show that the number of *open* bins with respect to the comparison chain grows only *sublinearly* in the number of items. Then, the optimality of FF2 follows.

We can use Theorem 83 to construct a simple online algorithm, which is $1/3$-competitive for general item sizes if the items arrive in random order. In order to do this we subdivide the items online into *size*-items and *color*-items. The size-items are then packed using Dual-Next-Fit and the color-items are packed using FF2.

▶ **Corollary 84.** *There exists an online algorithm with random-order ratio* $1/3$ *for class-constrained bin covering with general item sizes.*

This also gives us a randomized algorithm for the offline case. To the best of our knowledge this is the first offline algorithm presented for this problem.

▶ **Corollary 85.** *There exists a randomized offline algorithm* A *with* $\mathrm{AR}(A) = 1/3$ *for class-constrained bin covering.*

5.2 Deferred Proofs

5.2.1 Analysis of the ColorSets-approach with respect to the Average Performance Ratio

▶ **Theorem 74.** *Let* $\mathcal{D}$ *denote the set of all distributions on* $[0,1] \times \mathbb{N}$*, and let* $\epsilon > 0$ *be arbitrary. Choosing* M *sufficiently large, it holds*

$$\mathrm{APR}(\mathrm{CS}\left[\mathrm{HARMONIC}_M\right], \mathcal{D}) \leq h_\infty + \epsilon.$$

Furthermore, let $\epsilon > 0$ *be arbitrary. Then, there exists a parameter* k *for class-constrained bin packing such that for every online algorithm* A*, which is intended for classical bin packing, it holds*

$$\mathrm{APR}(\mathrm{CS}\left[A\right], \mathcal{D}) \geq h_\infty - \epsilon.$$

Proof of Theorem 74. At first we want to study the performance of $\mathrm{CS}\left[\mathrm{HARMONIC}_M\right]$. Let F be an arbitrary distribution on $[0,1] \times \mathbb{N}$. We can assume without loss of generality that the size of a drawn item is greater than zero with positive probability, otherwise the result of the algorithm is optimal.

At the beginning we show that $\mathrm{CS}\left[\mathrm{HARMONIC}_M\right]$ has a bounded competitive ratio so that we can use Lemma 31 afterwards. Assume we are given an instance I that contains items of q different colors. Let $S(I)$ denote the total size of the items in I, C_i the number of closed bins in the i-th group and O_i the number of open bins in the i-th group opened by the algorithm.

The number of groups is equal to $\lceil q/k \rceil$. Furthermore, the bin level of each closed bin is greater than $1/2$. Therefore, we obtain

$$\begin{aligned} \frac{\mathrm{CS}\left[\mathrm{HARMONIC}_M\right](I)}{\mathrm{OPT}(I)} &= \frac{\sum_{i=1}^{\lceil q/k \rceil}(O_i + C_i)}{\mathrm{OPT}(I)} \\ &\leq \frac{M\lceil q/k \rceil + \sum_{i=1}^{\lceil q/k \rceil} C_i}{\mathrm{OPT}(I)} \\ &\leq \frac{M\lceil q/k \rceil + 2S(I)}{\mathrm{OPT}(I)}. \end{aligned} \tag{23}$$

There are two obvious lower bounds for $\mathrm{OPT}(I)$:

$$\mathrm{OPT}(I) \geq S(I) \quad \text{and} \quad \mathrm{OPT}(I) \geq \lceil q/k \rceil. \tag{24}$$

Combining (23) and (24) it follows that

$$\frac{\mathrm{CS}\left[\mathrm{HARMONIC}_M\right](I)}{\mathrm{OPT}(I)} \leq \min\left\{ \frac{M\lceil q/k \rceil + 2S(I)}{\lceil q/k \rceil}, \frac{M\lceil q/k \rceil + 2S(I)}{S(I)} \right\}.$$

The last expression is maximized if $S(I) = \lceil q/k \rceil$, and therefore we obtain an upper bound of $M+2$ for the competitive ratio of the algorithm.

The next thing we need for the proof of the upper bound is an important observation about the number of distinct samples if the underlying distribution is realized on a countable set.

▶ **Lemma 86** (See Example 3.9 in [13]). *Let p be a probability mass function on $\mathbb{N}$. Let $X_1, \ldots, X_n$ be independent random variables that are distributed according to p. Let $Z_n = f(X_1, \ldots, X_n)$ denote the number of distinct values taken by these n random variables. Then we have*

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[Z_n\right]}{n} = 0.$$

Let Z_n denote the number of different drawn colors. Then, the number of groups generated by the algorithm is $\lceil Z_n/k \rceil$. It follows that

$$\begin{aligned}\mathbb{E}\left[\mathrm{CS}\left[\mathrm{HARMONIC}_M\right](I_n^F)\right] &\leq \mathbb{E}\left[\sum_{i=1}^{\lceil Z_n/k \rceil} (C_i + O_i)\right] \\ &\leq \mathbb{E}\left[\sum_{i=1}^{\lceil Z_n/k \rceil} C_i + M \cdot Z_n\right] \leq \mathbb{E}\left[\sum_{i=1}^{\lceil Z_n/k \rceil} C_i\right] + M\mathbb{E}\left[Z_n\right].\end{aligned}$$

For I_n^F let $\bar{I}_n^F$ denote the corresponding instance for classical bin packing that is induced by ignoring the colors and let $\mathrm{OPT}(\bar{I}_n^F)$ denote the value of the corresponding optimal solution. Then we obtain

$$\frac{\mathbb{E}\left[\sum_{i=1}^{\lceil Z_n/k \rceil} C_i\right] + \mathbb{E}\left[Z_n\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \leq \frac{\mathbb{E}\left[\mathrm{HARMONIC}_M(\bar{I}_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(\bar{I}_n^F)\right]} + \frac{M\mathbb{E}\left[Z_n\right]}{\mathbb{E}\left[S(I_n^F)\right]}.$$

It follows from the worst-case analysis of $\mathrm{HARMONIC}_M$ that we can find for $\epsilon > 0$ arbitrary an M sufficiently large, such that $\mathbb{E}\left[\mathrm{HARMONIC}_M(\bar{I}_n^F)\right] \leq (h_\infty + \epsilon)\mathbb{E}\left[\mathrm{OPT}(\bar{I}_n^F)\right]$ for n large enough. Furthermore, using Lemma 86 we see that $M\mathbb{E}\left[Z_n\right]/\mathbb{E}\left[S(I_n^F)\right]$ converges to zero as n tends to infinity. This yields the upper bound.

Now we show the lower bound for algorithms that are based on the ColorSets-approach. The proof is based on the famous work of C.C. Lee and D.T. Lee [58]. They established a lower bound for bounded-space online-algorithms for bin packing by grouping items in intervals according to their size. Then, bounded-space algorithms are forced to pack – more or less – all items from the same group together. We use here the same approach and show that we can find a distribution F such that with high probability items of the same size belong to the same group.

Let $t_1 = 1$ and $t_i = t_{i-1}(t_{i-1}+1)$. Then, we have $\sum_{i=1}^{\infty} \frac{1}{t_i+1} = 1$ and $1.691 \approx h_\infty = \sum_{i=1}^{\infty} \frac{1}{t_i}$. For technical details we refer to [58].

Let $\epsilon > 0$ and A an arbitrary online algorithm for classical bin packing. We want to show that we can find a parameter k and a distribution F, which depends on k, such that the average performance ratio of $\mathrm{CS}\left[A\right]$ is greater than $h_\infty - \epsilon$.

Now we specify F: We set $\gamma_k := 1 - \sum_{i=1}^{k} \frac{1}{t_i+1}$. The multiset of items $\mathcal{I}$ is as follows: The items have colors $1, \ldots, k^2$. We distinguish between *large* and *small* items. There are k^2 large items: For each $(i,j) \in [k]^2$ we have an item of size $\frac{1}{t_i+1} + \frac{\gamma_k}{2k}$ and color $(i-1)k + j$. Furthermore, the small items are as follows: For each color $i \in [k^2]$ we have $m_i \in \mathbb{N}$ items of size $\frac{\gamma_k}{2km_i}$. The multiplicities m_i will be specified later.

We note that for $i \in [k]$ we can group the colors $i + (j-1)k$ with $j \in [k]$ so that all items of these colors fit perfectly into a bin, that is the total size of the corresponding items in $\mathcal{I}$ is equal to 1. We want to show that we can find values $m_1, \ldots, m_{k^2}$ such that with high probability the colors appear in the order $1, \ldots, k^2$. Then, the algorithm CS $[A]$ packs large items of the same size together.

In order to do this let T_n^i denote the point in time when in I_n^F the first item of color i arrives. Let $\mathcal{E}_{\text{order}}$ denote the event $\{T_n^1 < T_n^2 < \ldots < T_n^{k^2} \leq n\}$.

▶ **Lemma 87.** *Let $\epsilon > 0$ be arbitrary. We can find values $m_1, \ldots, m_{k^2}$ such that for n sufficiently large, we have $\mathbb{P}[\mathcal{E}_{\text{order}}] \geq 1 - \epsilon$.*

Proof. Using de Morgan's laws and a union bound we obtain

$$\begin{aligned}
&\mathbb{P}\left[T_n^1 < T_n^2 < \ldots < T_n^{k^2} \leq n\right] \\
=&\mathbb{P}\left[\{T_n^1 < T_n^2\} \cap \{T_n^2 < T_n^3\} \cap \ldots \cap \{T_n^{k^2-1} < T_n^{k^2}\} \cap \{T_n^{k^2} \leq n\}\right] \\
=&1 - \mathbb{P}\left[\{T_n^2 < T_n^1\} \cup \{T_n^3 < T_n^2\} \cup \ldots \cup \{T_n^{k^2} < T_n^{k^2-1}\} \cup \{T_n^{k^2} = \infty\}\right] \\
\geq&1 - \mathbb{P}\left[T_n^2 < T_n^1\right] - \mathbb{P}\left[T_n^3 < T_n^2\right] - \ldots - \mathbb{P}\left[T_n^{k^2} < T_n^{k^2-1}\right] - \mathbb{P}\left[T_n^{k^2} = \infty\right].
\end{aligned}$$

Let p_i denote the probability that we draw an item of color i. Then we have $p_i = (m_i + 1)/(k^2 + \sum_{i=1}^{k^2} m_i)$. It follows that

$$\begin{aligned}
\mathbb{P}\left[T_n^{j+1} < T_n^j\right] &= \sum_{u=1}^{n} (1 - p_j - p_{j+1})^{u-1} p_{j+1} \\
&= p_{j+1} \sum_{u=0}^{n-1} (1 - p_j - p_{j+1})^u \\
&= \frac{p_{j+1}}{p_j + p_{j+1}} - \frac{(1 - p_j - p_{j+1})^n}{p_j + p_{j+1}} \\
&= \frac{m_{j+1} + 1}{m_j + m_{j+1} + 2} - \frac{(1 - p_j - p_{j+1})^n}{p_j - p_{j+1}}.
\end{aligned}$$

Furthermore, we have $\mathbb{P}\left[T_n^{k^2} = \infty\right] = (1 - p_{k^2})^n$. Hence, we see that for arbitrary $\epsilon > 0$ if we choose $m_1 \gg m_2 \gg \ldots \gg m_{k^2}$, we have for n large enough $\mathbb{P}[\mathcal{E}_{\text{order}}] \geq 1 - \epsilon$. ◀

Finally, we investigate the performance of CS $[A]$ (I_n^F) and $\text{OPT}(I_n^F)$. In order to do this, we have to show that with high probability the number of drawn items of each type is roughly the expected number of such items. Let N_n^i denote the number of drawn large items of color i in I_n^F. We set $N_n^{\min} = \min\{N_n^1, \ldots, N_n^{k^2}\}$ and $N_n^{\max} = \max\{N_n^1, \ldots, N_n^{k^2}\}$. Moreover, let V_n^i denote the number of drawn small items of color i in I_n^F.

Let $\mathcal{E}_A$ denote the event

$$\left\{\frac{n}{k^2 + \sum_{i=1}^{k^2} m_i} - \sqrt{n\log(n)} \leq N_n^{\min} \leq N_n^{\max} \leq \frac{n}{k^2 + \sum_{i=1}^{k^2} m_i} + \sqrt{n\log(n)}\right\}.$$

Using a union bound and applying Hoeffding's inequality we obtain

$$\mathbb{P}[\mathcal{E}_A] \geq 1 - \frac{2k^2}{n^2}.$$

Furthermore, let

$$\mathcal{E}_B = \left\{ \forall\, i \in [k^2] \,:\, V_n^i \leq n \cdot \frac{m_i}{k^2 + \sum_{i=1}^{k^2} m_i} + m_i \sqrt{n \log(n)} \right\}.$$

Again, using a union bound and Hoeffding's inequality we see that

$$\mathbb{P}\left[\mathcal{E}_B\right] \geq 1 - \frac{k^2}{n^2}.$$

It follows that for arbitrary k and $\epsilon > 0$ we can find $m_1 \gg m_2 \gg \ldots \gg m_{k^2}$ such that for n large enough we have $\mathbb{P}\left[\mathcal{E}_{\text{order}} \cap \mathcal{E}_A \cap \mathcal{E}_B\right] \geq 1 - \epsilon$.

We begin estimating $\text{CS}\,[A]\,(I_n^F)$. If $\mathcal{E}_{\text{order}}$ takes place, then $\text{CS}\,[A]$ packs all large items of the same size together. Therefore, in this case we have

$$\text{CS}\,[A]\,(I_n^F) \geq \sum_{i=1}^{k} \left[\frac{1}{t_i} \sum_{j=1}^{k} N_n^{(i-1)k+j} \right] \geq k N_n^{\min} \sum_{i=1}^{k} \frac{1}{t_i}.$$

Moreover, if $\mathcal{E}_A \cap \mathcal{E}_B$ takes place, we can arrange the items in at most

$$n \cdot \frac{k}{k^2 + \sum_{i=1}^{k^2} m_i} + \sqrt{n \log(n)}$$

many bins using the perfect-packing induced by $\mathcal{I}$. Hence, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{\text{CS}\,[A]\,(I_n^F)}{\text{OPT}(I_n^F)}\right] &\geq \mathbb{E}\left[\frac{\text{CS}\,[A]\,(I_n^F)}{\text{OPT}(I_n^F)} \cdot I_{\{\mathcal{E}_{\text{order}} \cap \mathcal{E}_A \cap \mathcal{E}_B\}}\right] \\
&\geq \mathbb{E}\left[\frac{k N_n^{\min} \sum_{i=1}^{k} \frac{1}{t_i}}{n \cdot \frac{m_i}{k^2 + \sum_{i=1}^{k^2} m_i} + m_i \sqrt{n \log(n)}} \cdot I_{\{\mathcal{E}_{\text{order}} \cap \mathcal{E}_A \cap \mathcal{E}_B\}}\right] \\
&\geq \mathbb{E}\left[\frac{k \cdot \left(\frac{n}{k^2 + \sum_{i=1}^{k^2} m_i} - \sqrt{n \log(n)}\right) \cdot \sum_{i=1}^{k} \frac{1}{t_i}}{n \cdot \frac{k}{k^2 + \sum_{i=1}^{k^2} m_i} + \sqrt{n \log(n)}} \cdot I_{\{\mathcal{E}_{\text{order}} \cap \mathcal{E}_A \cap \mathcal{E}_B\}}\right] \\
&\geq \frac{k \cdot \left(\frac{n}{k^2 + \sum_{i=1}^{k^2} m_i} - \sqrt{n \log(n)}\right) \cdot \sum_{i=1}^{k} \frac{1}{t_i}}{n \cdot \frac{k}{k^2 + \sum_{i=1}^{k^2} m_i} + \sqrt{n \log(n)}} \cdot (1 - \epsilon).
\end{aligned}$$

Choosing k large enough we have $\sum_{i=1}^{k} \frac{1}{t_i} \geq h_\infty - \epsilon$. The statement follows as n tends to infinity. ◀

5.2.2 Lower Bounds for Selected Heuristics for Class-constrained Bin Packing with respect to the Random-order Ratio

▶ **Proposition 75.** *Let A be an arbitrary online algorithm for classical bin packing. If $k = 2$, it holds $\text{RR}(\text{CS}\,[A]) \geq 2$. Moreover, for every ϵ greater than zero, there exists a parameter k such that $\text{RR}(\text{FF}) \geq 2 - \epsilon$.*

Proof of Proposition 75. We start by showing the lower bound for algorithms based on the ColorSets-approach. Let $k = 2$ and the instance I be as follows: The items in I can be packed in N bins. Let $\gamma > 0$ be sufficiently small. The i-th bin contains items of the two colors $2i-1$ and $2i$. There are m_{2i-1} items of size γ/m_{2i-1} and one item of size $1/2 - (i+1)\gamma$ of color $2i-1$ and m_{2i} many items of size γ/m_{2i} and one item of size $1/2 + (i-1)\gamma$ of color $2i$. We denote by T_i with $1 \leq i \leq 2N$ the first time an item of color i arrives. Let $\mathcal{E}$ denote the event

$$\mathcal{E} = \{T_1 < T_4 < T_3 < T_6 < T_5 < T_8 < \ldots < T_{2N-3} < T_{2N}\}.$$

If $\mathcal{E}$ takes place the total size of the items in all but the last color set is greater than 1. Therefore, CS $[A]$ has to open at least $2N-1$ many bins.

It remains to show that for $\epsilon > 0$ we can choose I in such a way that $\mathbb{P}[\mathcal{E}] \geq 1-\epsilon$. We note that there are $m_i + 1$ many items of color i and that we have for two colors i, j with $i \neq j$

$$\mathbb{P}[T_i < T_j] = \frac{m_i + 1}{(m_i+1) + (m_j+1)}.$$

Therefore, we obtain

$$\begin{aligned}
&\mathbb{P}[T_1 < T_4 < T_3 < T_6 < \ldots < T_{2N-3} < T_{2N}] \\
=&\mathbb{P}[\{T_1 < T_4\} \cap \{T_4 < T_3\} \cap \{T_3 < T_6\} \cap \ldots \cap \{T_{2N-3} < T_{2N}\}] \\
=&1 - \mathbb{P}[\{T_4 < T_1\} \cup \{T_3 < T_4\} \cup \{T_6 < T_3\} \cup \ldots \cup \{T_{2N} < T_{2N-3}\}] \\
\geq&1 - (\mathbb{P}[T_4 < T_1] + \mathbb{P}[T_3 < T_4] + \mathbb{P}[T_6 < T_3] + \ldots + \mathbb{P}[T_{2N} < T_{2N-3}]) \\
=&1 - \left(\frac{m_4+1}{m_1+m_4+2} + \frac{m_3+1}{m_3+m_4+2} + \ldots + \frac{m_{2N}+1}{m_{2N-3}+m_{2N}+1}\right).
\end{aligned}$$

We see that if we choose the m_i in such a way that $m_1 \gg m_4 \gg m_3 \gg m_6 \gg \ldots \gg m_{2N-3} \gg m_{2N}$, we obtain $\mathbb{P}[\mathcal{E}] \geq 1-\epsilon$.

It follows that

$$\frac{\mathbb{E}[\mathrm{CS}[A](I^\sigma)]}{\mathrm{OPT}(I)} \geq (1-\epsilon)\frac{2N-1}{N}.$$

To show the lower bound for First-Fit we proceed in the following way: Let $N \in \mathbb{N}$, with N a multiple of k, and $\gamma \in (0, \frac{1}{2(k-1)})$. The items in I can be arranged in N perfectly packed bins. Here, the i-th bin contains m items of size γ/m of each of the colors $(i-1)k+1, \ldots, ik-1$ and one item of size $1-(k-1)\gamma$ of color ik. We call the colors j with $(j \mod k) \neq 0$ *small* and otherwise *large*. Let $\mathcal{E}$ denote the event that for each small color an item arrived before the first item of a large color arrives. Let $\epsilon > 0$ be arbitrary. Arguing as in the first part we observe that choosing m large enough we can enforce that $\mathbb{P}[\mathcal{E}] \geq 1-\epsilon$. If $\mathcal{E}$ takes place First-Fit at first opens $N(k-1)/k$ many bins for the small colors. Afterwards, we have to open a new bin for each item of a large color. Therefore, First-Fit has to open $N(k-1)/k + N = (2-1/k)N$ many bins. Hence, $\mathbb{E}[\mathrm{FF}(I^\sigma)] / \mathrm{OPT}(I) \geq (1-\epsilon)\cdot(2-1/k)$. ◀

5.2.3 Analysis of Selected Heuristics for Class-constrained Bin Packing with Unit Sized Items with respect to the Average Performance Ratio

▶ **Proposition 76.** *We consider class-constrained bin packing with unit sized items with parameters B and k. Let $\mathcal{D}$ denote the set of all possible distributions, and let A be* CS [NF] *or any online algorithm that opens a new bin only if it is forced, then it holds*

$$\mathrm{APR}(A, \mathcal{D}) = 1.$$

Proof of Proposition 76. Let F be an arbitrary distribution and A be an arbitrary algorithm, satisfying the conditions of the statement. We say that a bin is *open* if the number of contained items is smaller than B, otherwise we call it *closed*. We introduce the following random variables: Z_n denotes the number of distinct colors in I_n^F, C_n^A denotes the number of closed and O_n^A the number of open bins of A, after I_n^F has been processed. We observe that O_n^A is upper bounded by Z_n. Using the estimate $\mathrm{OPT}(I_n^F) \geq n/B$, we have

$$\frac{\mathbb{E}\left[A(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \leq \frac{B\left(\mathbb{E}\left[C_n^A\right] + \mathbb{E}\left[O_n^A\right]\right)}{n}$$
$$\leq \frac{B \cdot \mathbb{E}\left[C_n^A\right]}{n} + \frac{B \cdot \mathbb{E}\left[Z_n\right]}{n} \leq \frac{B \cdot \lceil n/B \rceil}{n} + B \cdot \frac{\mathbb{E}\left[Z_n\right]}{n}.$$

Since $\mathbb{E}\left[Z_n\right]/n$ tends to zero as n tends to infinity (see Lemma 86), the previous expression converges to 1 as n tends to infinity. This shows the optimality. ◀

5.2.4 A Lower Bound for the Random-order Ratio of $\mathrm{CS}\left[\mathrm{NF}\right]$ in the case of Unit Sized Items

▶ **Proposition 77.** *For every ϵ greater than zero there exist parameters B and k for class-constrained bin packing with unit sized items such that* $\mathrm{RR}(\mathrm{CS}\left[\mathrm{NF}\right]) \geq 2 - \epsilon$.

Proof of Proposition 77. The random-order ratio is upper bounded by 2, because this is the competitive ratio of the algorithm [77]. Now we show a lower bound, which matches this upper bound. For $B, k \in \mathbb{N}$, we will look at the following instance: There are a bins, where we can choose a arbitrary. The i-th bin contains $B-(k-1)$ items of color 1, and one of each of the colors $\{(i-1)\cdot(k-1)+2, \ldots, i\cdot(k-1)+1\}$. With probability $\frac{B-(k-1)}{B} = 1 - \frac{k-1}{B} =: 1-\epsilon$ the first drawn item has color 1. Then, the first group contains color 1, and $k-1$ items of small colors.

In the first group, there are $a(B+1-k)+k-1$ many items. So we need $\lceil \frac{a(B+1-k)+(k-1)}{B} \rceil$ bins to pack this group. Furthermore, there are $(a-1)\cdot(k-1)$ many colors left. So there are additional $\lceil \frac{(a-1)\cdot(k-1)}{k} \rceil$ many groups. We need only one bin for each additional group. Hence, the expected number of used bins is lower bounded by

$$(1-\epsilon)\cdot\left(\frac{a(B+1-k)+(k-1)}{B} + \frac{(a-1)\cdot(k-1)}{k} + 2\right) + \epsilon \cdot 2a.$$

Divided by OPT, i.e., a, we achieve

$$(1-\epsilon)\cdot\left(1-\epsilon+\frac{\epsilon}{a} + \left(1-\frac{1}{a}\right)\cdot\left(1-\frac{1}{k}\right) + \frac{2}{a}\right) + 2\epsilon.$$

Hence, if we choose at first k, then B, and finally a large enough, the random-order ratio of the ColorSets-algorithm is arbitrary close to 2. ◀

5.2.5 A Lower Bound for the Random-order Ratio of First-Fit in the case of Unit Sized Items

▶ **Proposition 78.** *For every ϵ greater than zero there exist parameters B and k for class-constrained bin packing with unit sized items such that* $\mathrm{RR}(\mathrm{FF}) \geq 1.5 - \epsilon$.

Proof of Proposition 78. The proof of the statement is based on analyzing the following set of items: Let $a, m \in \mathbb{N}_{\geq 2}$. We set $k = 2$ and $B = m^2$. The items in $\mathcal{I}$ can be packed in a

bins, where the i-th bin contains $m^2 - 1$ items of color i and one item of color $a+i$. We call the colors $\{1, \ldots, a\}$ *large* and $\{a+1, \ldots, 2a\}$ *small*.

Let $\mathcal{B}^\sigma$ denote the set of bins that were opened by First-Fit after I^σ has been processed. Let $\mathcal{B}_1^\sigma \subseteq \mathcal{B}^\sigma$ denote the set of bins, containing items of two different colors, which were opened by one of the first ma arriving items. Furthermore, let $\mathcal{B}_2^\sigma = \mathcal{B}^\sigma \backslash \mathcal{B}_1^\sigma$. Moreover, let $\mathcal{C}^\sigma$ denote the set of colors, for which at least one item of this color is put into a bin from $\mathcal{B}_2^\sigma$. Then, we obtain

$$\mathrm{FF}(I^\sigma) = |\mathcal{B}_1^\sigma| + |\mathcal{B}_2^\sigma| \geq |\mathcal{B}_1^\sigma| + |\mathcal{C}^\sigma|/2.$$

In the following we want to show that with high probability $|\mathcal{B}_1^\sigma| \geq (1/2 - \epsilon)a$ and $|\mathcal{C}^\sigma| \geq (2-\epsilon)a$.

In order to do this, we will introduce a couple of events and show that First-Fit will not be better than $(1.5 - \epsilon)$-competitive if all events occur simultaneously. Finally, we will show that all these events occur with high probability.

Let $SC(I^\sigma)$ denote the number of items of small colors among the first ma items drawn in I^σ, let $N(I^\sigma)$ denote the number of different drawn large colors among the first ma drawn items, and finally let $S(I^\sigma)$ denote the number of large colors, for which an item of these colors is put into a bin from $\mathcal{B}_2^\sigma$.

We introduce the following three events:

$$\mathcal{E}_1 = \left\{ SC(I^\sigma) \leq \frac{2}{m} \cdot a \right\},$$
$$\mathcal{E}_2 = \left\{ N(I^\sigma) \geq \left(1 - \frac{4}{m}\right) a \right\}, \quad \textit{and}$$
$$\mathcal{E}_3 = \left\{ S(I^\sigma) \geq \left(1 - \frac{6}{m}\right) a \right\}.$$

Assume all three events occur simultaneously: Then, it follows immediately that

$$|\mathcal{B}_1^\sigma| \geq \frac{1}{2}\left(1 - \frac{4}{m}\right) a = \left(\frac{1}{2} - \frac{2}{m}\right) a$$

and

$$|\mathcal{C}^\sigma| \geq \left(1 - \frac{2}{m}\right) a + \left(1 - \frac{6}{m}\right) a = \left(2 - \frac{8}{m}\right) a.$$

Therefore, we have $\mathrm{FF}(I^\sigma) \geq (3/2 - 6/m)a$. Choosing m large enough and then dividing by a yields a lower bound for the random-order ratio of $1.5 - \epsilon$ on $\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3$.

Now it remains to show that $\mathbb{P}[\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3]$ converges to 1 as a tends to infinity.

▶ **Lemma 88.** *It is*

$$\mathbb{P}[\mathcal{E}_1] \geq 1 - \exp\left(-\frac{2}{3m} \cdot a\right).$$

Proof. Let $X_i = I_{\{\text{color of the } i\text{-th arrived item in } I^\sigma \text{ is small}\}}$. Then, we have $\mathbb{E}[X_i] = 1/m^2$. Therefore,

$$\mathbb{P}[\mathcal{E}_1] = \mathbb{P}\left[\sum_{i=1}^{am} X_i \geq \frac{2}{m} \cdot a\right] = \mathbb{P}\left[\sum_{i=1}^{am} X_i - \sum_{i=1}^{am} \mathbb{E}[X_i] \geq \frac{a}{m}\right].$$

The random variables X_i are not independent. But we can apply the concentration bound from Proposition 13, since we can look at the X_i as samples drawn according to sampling without replacement from a finite population. Then the result follows immediately. ◀

▶ **Lemma 89.** *It is*

$$\mathbb{P}\left[\mathcal{E}_2 \mid \mathcal{E}_1\right] \geq 1 - \frac{1}{a^2}.$$

Proof. Since we know that $\mathcal{E}_1$ takes place we can assume without loss of generality that we draw at least $(m - 2/m)a$ many items of large colors. At first we look at the case where the items of large colors are drawn with respect to sampling with replacement. That is, we assume that the large colors are drawn according to the uniform distribution on $\{1, \dots, a\}$. Let D denote the number of different drawn colors if we draw $(m - 2/m)a$ items. It holds (see e.g. Example 3.9 in [13])

$$\mathbb{E}\left[D\right] = a \cdot \left[1 - \left(1 - \frac{1}{a}\right)^{(m-2/m)a}\right].$$

Using standard inequalities we can lower bound this expression by $(1 - 3/m)a$. The function, which expresses the number of different drawn large items satisfies the bounded difference property, so we can apply McDiarmid's inequality. This yields (if a is sufficiently large)

$$\mathbb{P}\left[D \geq \left(1 - \frac{4}{m}\right)a\right] \geq \mathbb{P}\left[D \geq \mathbb{E}\left[D\right] - \sqrt{ma\log(a)}\right] \geq 1 - \frac{1}{a^2}.$$

But we are in a setting, where the items are drawn according to sampling without replacement. To show that nonetheless the estimate is true, we will use a coupling argument. Assume the items with a large color are numbered from 1 to $a(m^2 - 1)$. We have two urns each of them containing all items of a large color. At first we generate an instance I by drawing $(m - 2/m)a$ many items according to sampling with replacement from the first urn. Now we want to generate an instance I'. Thereto, let U denote the set of indices of the drawn items in I. Now, we draw all items with an index in U from the second urn and remove them from the urn. In case that there are any duplicates, we also draw the missing $(m - 2/m)a - |U|$ items from the second urn according to sampling without replacement. This standard-procedure yields a coupling between sampling with and without replacement. That is, we have a pair of random variables (I, I') such that the marginals of this two-dimensional random variable are instances generated using sampling with and without replacement.

Let $\mathrm{DC}\left(I\right)$ denote the number of different drawn colors in I. Then, we have $\mathrm{DC}\left(I\right) \sim D$ and $\mathrm{DC}\left(I'\right) \sim N(I^\sigma)$. Now, using our coupling argument, it follows that

$$\begin{aligned}
\mathbb{P}\left[\mathrm{DC}\left(I\right) \geq (1 - 4/m)a\right] &= \sum_{\iota} \mathbb{P}\left[I = \iota\right] I_{\{\mathrm{DC}(\iota) \geq (1-4/m)a\}} \\
&= \sum_{\iota}\sum_{\iota'} \mathbb{P}\left[I = \iota, I' = \iota'\right] I_{\{\mathrm{DC}(\iota) \geq (1-4/m)a\}} \\
&= \sum_{\iota}\sum_{\iota'} \mathbb{P}\left[I' = \iota' \mid I = \iota\right] \mathbb{P}\left[I = \iota\right] I_{\{\mathrm{DC}(\iota) \geq (1-4/m)a\}} \\
&\leq \sum_{\iota}\sum_{\iota'} \mathbb{P}\left[I' = \iota' \mid I = \iota\right] \mathbb{P}\left[I = \iota\right] I_{\{\mathrm{DC}(\iota') \geq (1-4/m)a\}} \\
&= \mathbb{P}\left[\mathrm{DC}\left(I'\right) \geq (1 - 4/m)a\right].
\end{aligned}$$

Combining the statements yields the proof. ◀

▶ **Lemma 90.** *For each $m \in \mathbb{N}$ there exists a positive constant C_m such that*

$$\mathbb{P}\left[\mathcal{E}_3 \mid \mathcal{E}_2 \cap \mathcal{E}_1\right] \geq 1 - \exp\left(-C_m \cdot a\right).$$

Proof. We call a bin in $\mathcal{B}_1^\sigma$ special, if it contains items of two large colors, and no item of these two large colors is contained in a different bin in $\mathcal{B}_1^\sigma$. Since we assume that $\mathcal{E}_1 \cap \mathcal{E}_2$ occurs, we know that $|\mathcal{B}_1^\sigma| \geq (1/2 - 2/m)a$, and $SC(I^\sigma) \leq 2/m \cdot a$. Moreover, there can be at most $am/m^2 = a/m$ many closed bins when the $(am+1)$-th item arrives. Therefore, there have to be at least

$$\frac{\left(1 - \frac{4}{m} - \frac{2}{m} - \frac{2}{m}\right) a}{2} = \left(\frac{1}{2} - \frac{4}{m}\right) a$$

many special bins. Without loss of generality these bins are numbered from 1 to $(1/2 - 4/m)a$. We are now interested in the number of such bins, where both colors *survive*, that is they do not contain $B-1$ items of the same color. Let $Y_i = I_{\{\text{both colors in special bin } i \text{ survive}\}}$. We observe, that the random variables Y_i are independent, Bernoulli-distributed random variables.

Using the hypergeometric distribution and standard estimates, we see that

$$\mathbb{E}\left[Y_i\right] = 1 - \frac{2}{\frac{m^2-1}{m^2} \cdot \binom{2m^2-2}{m^2-1}}. \geq 1 - \frac{4}{\binom{2m^2-2}{m^2-1}} \geq 1 - \frac{1}{2^{m^2-3}} \geq 1 - \frac{1}{m}.$$

Therefore, using again Hoeffding's inequality, we see that we have at least

$$\left(\frac{1}{2} - \frac{4}{m}\right) \cdot \left(1 - \frac{2}{m}\right) a \geq \left(\frac{1}{2} - \frac{6}{m}\right) a$$

many specials bins, in which both colors survive, with a probability, which tends to 1 exponentially in a for arbitrary m. ◀

This proves Proposition 78. ◀

5.2.6 Lower Bounds for Bounded-space Algorithms for Class-constrained Bin Packing with Unit Sized Items

▶ **Proposition 79.** *Consider class-constrained bin packing with unit sized items and parameters B and k. Let A be an arbitrary bounded-space online algorithm. Then we have* $\mathrm{RR}(A), \mathrm{APR}(A, \mathcal{D}) \in \Omega(B/k)$.

Proof of Proposition 79. Let A be an ℓ-bounded space algorithm. Since every bin that A opens contains at least one item, A is at least B-competitive. Therefore, we can apply Lemma 31. So we look for a lower bound for the ratio $\mathbb{E}\left[A(I_n^F)\right] / \mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]$.

We show that the statement is even true if we allow the algorithm to use bins with infinite capacity. Let F be given by $\mathcal{I} = \{(1,1), \ldots, (1,m)\}$ and p is the uniform distribution on $\mathcal{I}$. We call an arriving item *nice* if no item of the same color is currently contained in any open bin of A. Then, we have $\mathbb{P}\left[A_i \text{ is nice}\right] \geq 1 - \frac{\ell k}{m}$. Let N denote the number of nice items. Then, a bounded space algorithm needs on average at least $\mathbb{E}\left[N\right]/k$ many bins to pack the items. Therefore, we have

$$\mathbb{E}\left[A(I_n^F)\right] \geq \frac{\mathbb{E}\left[N\right]}{k} \geq \frac{\left(1 - \frac{\ell k}{m}\right) n}{k}.$$

A valid packing is to pack each color separately using Next-Fit. It follows that

$$\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] \leq m + \frac{n}{B}.$$

Combining the previous estimates, we see that

$$\frac{\mathbb{E}\left[A(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \geq \frac{B}{k} \cdot \left(1 - \frac{\ell k}{m}\right) \cdot \frac{1}{1 + \frac{Bm}{n}}.$$

Since we can choose m arbitrary large, the statement follows. ◀

5.2.7 An Upper Bound for the Competitive Ratio in Class-constrained Bin Covering with Unit Sized Items

▶ **Proposition 80.** *The competitive ratio of any deterministic online algorithm for class-constrained bin covering with unit sized items is at most* $\left(H_{k-1} + 1 - \frac{k-1}{B}\right)^{-1}$. *If* $B = k$ *this yields an upper bound of* H_k^{-1}. *The same is true for randomized algorithms.*

Proof of Proposition 80. The proof follows along the lines of [37], but slightly changes the choice of the scenarios. Let N be an arbitrarily large number, divisible by $B!$. In the first phase, the algorithm receives N items of each of the colors $1, \ldots, k$. After this, a second phase follows. There are k different scenarios for the the second phase: In the zero-th scenario no more items will arrive. In the i-th scenario ($1 \leq i \leq k-1$) there arrive sufficiently many items of the colors $k+1$ up to $k+i$. The optimal solution in the zero-th scenario covers Nk/B bins, and $Nk/(k-i)$ bins in the i-th scenario.

Without loss of generality we can assume that an arbitrary but fixed online algorithm, uses only bins of the following type after the end of the first phase: bins with B items and k different colors (type 0), and bins with $k-i$ many items, where all colors are different (type i). We denote the number of bins of type i as x_i. We have

$$Bx_0 + \sum_{j=1}^{k-1} jx_j = kN. \tag{25}$$

To be R-competitive, the following conditions must be satisfied:

$$x_0 \geq R \cdot \frac{Nk}{B}, \tag{26}$$

$$x_0 + \sum_{j=k-i}^{k-1} x_j \geq R \cdot \frac{Nk}{k-i} \quad \forall\, i = 1, \ldots, k-1, \tag{27}$$

$$x_0, \ldots, x_{k-1} \geq 0.$$

Now we sum up all the constraints in (27) and $(B+1-k)$ times constraint (26). Combining this with (25) we obtain

$$Nk = \sum_{i=1}^{k-1} \left(x_0 + \sum_{j=k-i}^{k-1} x_j \right) + (B+1-k)x_0 \geq R \cdot \left(Nk \cdot H_{k-1} + Nk \cdot \frac{B-k+1}{B} \right).$$

Therefore, $R \leq \left(H_{k-1} + 1 - \frac{k-1}{B}\right)^{-1}$.

The statement is also true for randomized algorithms: The proof is based on Yao's principle [85]. Assume that the zero-th scenario occurs with probability $1 - \frac{k-1}{B} = \frac{B-k+1}{B}$, and the i-th scenario occurs with probability $\frac{1}{B}$, where $1 \leq i \leq k-1$. Then, we have

$$\mathbb{E}\left[\mathrm{OPT}\right] = \frac{B-k+1}{B} \cdot \frac{Nk}{B} + \frac{1}{B} \sum_{i=1}^{k-1} \frac{Nk}{k-i}$$

$$= \frac{B-k+1}{B} \cdot \frac{Nk}{B} + \frac{NK}{B} H_{k-1}$$
$$= \frac{Nk}{B} \cdot \left(1 + H_{k-1} - \frac{k-1}{B}\right).$$

We can show as in the deterministic case that

$$Nk = \sum_{i=1}^{k-1} \left(x_0 + \sum_{j=k-i}^{k-1} \mathbb{E}\left[x_j\right] \right) + (B+1-k)\mathbb{E}\left[x_0\right]$$
$$\geq R \cdot B \cdot \mathbb{E}\left[\mathrm{OPT}\right]$$
$$= R \cdot B \cdot \frac{Nk}{B} \cdot \left(1 + H_{k-1} - \frac{k-1}{B}\right).$$

It follows that

$$R \geq \left(1 + H_{k-1} - \frac{k-1}{B}\right)^{-1}.$$

Therefore, we obtain the same lower bound for randomized algorithms as for deterministic algorithms. ◀

5.2.8 An Upper Bound for the Competitive Ratio of Bounded-space Algorithms for Class-constrained Bin Covering

▶ **Proposition 81.** *Let A be a bounded-space algorithm for class-constrained bin covering. Then, it holds* $\mathrm{CR}(A) = 0$*.*

Proof of Proposition 81. Let k and B be arbitrary parameters for the problem, and let A be an ℓ-bounded space algorithm. Consider the following sequence of unit sized items: We start with $a \cdot (B+1-k)$ many items of color 1. Then, a items of each of the colors $\{2, \ldots, k\}$ arrive. An ℓ-bounded space algorithm can cover at most ℓ bins, but the optimal solution can cover a bins. ◀

5.2.9 Analysis of Dual-Next-Fit with respect to Probabilistic Performance Measures

▶ **Theorem 82.** *For class-constrained bin covering with unit sized items and parameters k and B it holds* $\mathrm{RR}(\mathrm{DNF}) \in \Theta(\log(k)^{-1})$*. In the case of general item sizes let $\mathcal{D}$ denote the set of all discrete perfect-packing distributions. Then we have* $\mathrm{APR}(\mathrm{DNF}, \mathcal{D}) \in \Theta(\log(k)^{-1})$*.*

Proof of Theorem 82. We can describe the state of Dual-Next-Fit as a pair (u, c), where $u \in \mathbb{R}_{\geq 0}$ denotes the total size of items, and $c \subseteq \mathbb{N}$ the set of colors in the bin currently processed. We can partition the possible states into the set of open bin configurations $\mathcal{O}$ and the set of covered bin configurations $\mathcal{C}$. We have

$$\mathcal{O} = \{(u, c) \,:\, (u < 1) \vee (|c| < k)\} \quad \text{and}$$
$$\mathcal{C} = \{(u, c) \,:\, (u \geq 1) \wedge (|c| \geq k)\}.$$

We can subsume all states in $\mathcal{C}$ to a single state $*$. Furthermore, we introduce another notion: $\mathrm{DNF}(s, I)$ denotes the number of covered bins performing on I if Dual-Next-Fit starts with the open bin configuration s.

▶ **Observation 91.** *Dual-Next-Fit satisfies the following two properties:*

- *Let I be an instance, and let $\tilde{I}$ be an instance that we obtain from I by decreasing the size of an item or deleting it. Then, we have* $\mathrm{DNF}(\tilde{I}) \leq \mathrm{DNF}(I)$.
- *Let $s, s' \in \mathcal{O} \cup \{*\}$ and I be arbitrary. Then, we have*

$$|\,\mathrm{DNF}(s, I) - \mathrm{DNF}(s', I)| \leq 1.$$

We begin with considering the case that the items are drawn independently and identically distributed according to a discrete perfect-packing distribution F. Let $\mathcal{I}$ be the corresponding set, and $a, b \in \mathbb{N}$ such that $\mathrm{OPT}(\mathcal{I}) = a$ and $|\mathcal{I}| = ak + b$. Since we are dealing with a maximization problem, we can apply Lemma 31. Therefore, it suffices to give bounds for $\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right]$ and $\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]$.

At first we give an upper bound for $\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]$. Let $S(I_n^F)$ denote the total size of the items drawn in I_n^F. Then, we have

$$\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right] \leq \mathbb{E}\left[S(I_n^F)\right] = n \cdot \frac{a}{ak+b}. \tag{28}$$

Now, we give a lower bound for $\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right]$. For this purpose we use the tools from the field of Markov chains. We can view the behavior of Dual-Next-Fit as a Markov chain (X_n) on the state space $S \subseteq \mathcal{O} \cup \{*\}$. ($S$ is given by all states, which can be reached with positive probability.) Except for the case $k = 1$ this yields an irreducible and aperiodic Markov chain. An irreducible and aperiodic Markov chain possesses a unique *stationary distribution* π^F on S, and the distribution of X_n converges to π^F. Then, we have

$$\mathbb{E}_{s \sim \pi^F}\left[\mathrm{DNF}(s, I_n^F)\right] = n \cdot \pi^F(*).$$

We know from Proposition 21 that it holds $\pi^F(*) = \mathbb{E}\left[T\right]^{-1}$, where T denotes the first return time to $*$ starting in $*$. Hence, it follows from Observation 91 that

$$\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right] \geq \mathbb{E}_{s \sim \pi^F}\left[\mathrm{DNF}(s, I_n^F)\right] - 1 = \frac{n}{\mathbb{E}\left[T\right]} - 1. \tag{29}$$

The interesting thing is that we can estimate $\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right]$ in terms of $\mathbb{E}\left[T\right]$.

Now we will estimate the return-time T. We remember that F is induced by a multiset $\mathcal{I}$ with $\mathrm{OPT}(\mathcal{I}) = a$ and $|\mathcal{I}| = ak + b$. In order to upper bound T we look at the number of items we need to fulfill the color- and the size-condition independently. Let T_s denote the number of items we need until their total size is at least 1. Moreover, let T_c denote the number of items until we have drawn k different colors (starting with no drawn color). Then,

$$\mathbb{E}\left[T\right] \leq \mathbb{E}\left[T_s\right] + \mathbb{E}\left[T_c\right]. \tag{30}$$

We have

$$\mathbb{E}\left[T_s\right] \leq 2 \cdot \frac{ak+b}{a}, \tag{31}$$

otherwise Dual-Next-Fit would not be a 1/2-competitive algorithm for classical bin covering in the worst case. So it remains to give an upper bound for $\mathbb{E}\left[T_c\right]$. For this purpose, we choose k items of different colors for each of the a bins inducing F. The unconsidered items are removed from the bins. We denote the uniform distribution on the remaining items by $\tilde{F}$. Then, we have $\mathbb{E}\left[T_c\right] \leq \frac{ak+b}{ak} \cdot \mathbb{E}\left[\tilde{T}_c\right]$. $\mathbb{E}\left[\tilde{T}_c\right]$ is maximized if there are a items of each of the colors 1 to k. (This follows from the fact that for each considered distribution, we have probability at least $(k-i)/k$ to draw a missing color if there are already i different colors drawn.) But in this case it follows from the Coupon-collectors problem that we can upper bound $\mathbb{E}\left[\tilde{T}_c\right]$ by kH_k. Hence, combining (28), (30) and (31) we obtain

$$\mathbb{E}[T] \leq (2+H_k) \cdot \frac{ak+b}{a}.$$

Then, combining (28), (29) and the previous estimate we see that

$$\frac{\mathbb{E}\left[\mathrm{DNF}(I_n^F)\right]}{\mathbb{E}\left[\mathrm{OPT}(I_n^F)\right]} \geq \frac{1}{2+H_k} - \frac{1}{n} \cdot \frac{ak+b}{a}. \tag{32}$$

This gives us a lower bound for the performance of Dual-Next-Fit in the case the items are drawn independently and identically distributed. The bound is essentially tight: Think of the case where the items are drawn according to the uniform distribution from $\mathcal{I} = \{(1/k, 1), \ldots, (1/k, k)\}$.

Now, we will investigate the setting of class-constrained bin covering with unit sized items and parameters k and B, where the items from a multiset $\mathcal{I}$ arrive in random order. Because of the monotonicity of Dual-Next-Fit we can assume again that we can arrange the items in $\mathcal{I}$ perfectly in bins, that is we have $\mathrm{OPT}(\mathcal{I}) = |\mathcal{I}|/B$. We will reduce this setting to the case where unit sized items are drawn independently and identically distributed. The idea is to split up the instance into smaller subinstances. In a special case this was done in [20]. In [42] this idea was combined with an estimate based on the total variation distance. [60] covers the topic of total variation distance in Chapter 4.

▶ **Definition 92.** Let Ω be a discrete set. The total variation distance between two probability distributions μ and ν on Ω is defined by

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subseteq \Omega} |\mu(A) - \nu(A)|.$$

One can show that $\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2}\sum_{x \in \Omega} |\mu(x) - \nu(x)|$.

▶ **Lemma 93.** *Let μ, ν be two probability distributions on Ω. Let $X, Y : \Omega \to \mathbb{R}$ be random variables with $X \sim \mu$ and $Y \sim \mu$. Furthermore, let $f : \mathbb{R} \to \mathbb{R}$. Then we obtain*

$$|\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]| \leq 2\|\mu - \nu\|_{\mathrm{TV}} \cdot \|f\|_\infty.$$

Proof. It is

$$\begin{aligned} |\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]| &= \left|\sum_{x \in \Omega} (\mu(x) - \nu(x)) \cdot f(x)\right| \\ &\leq 2\|f\|_\infty \cdot \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = 2\|f\|_\infty \cdot \|\mu - \nu\|_{\mathrm{TV}}. \end{aligned}$$

◀

We set $m = \sup\{u \in \mathbb{N} : u^3 \leq |I|\}$.

▶ **Lemma 94.** *Let F be the discrete perfect-packing distribution induced by $\mathcal{I}$, that is the uniform distribution on $\mathcal{I}$. Then, we have*

$$\mathbb{E}[\mathrm{DNF}(I^\sigma)] \geq m^2 \mathbb{E}\left[\mathrm{DNF}(I_m^F)\right] - \mathcal{O}(m^2).$$

Proof. Let $I^\sigma = (A_{\sigma(1)}, \ldots A_{\sigma(|I|)})$ denote a random permutation of our given instance. Let $\tilde{I}^\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(m^3)})$ denote the truncated instance. Since $\mathbb{E}\left[\mathrm{DNF}(I^\sigma)\right] \geq \mathbb{E}\left[\mathrm{DNF}(\tilde{I}^\sigma)\right]$ we will consider the truncated instance in the following.

We want to partition $\tilde{I}^\sigma$ into m^2 subinstances of length m. In order to do this, we set $\tilde{I}_1^\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(m)})$, $\tilde{I}_2^\sigma = (A_{\sigma(m+1)}, \ldots, A_{\sigma(2m)})$, and so on. We see that the random variables $\tilde{I}_j^\sigma$ are identically distributed. Hence, from Observation 91 it follows that

$$\begin{aligned}\mathbb{E}\left[\mathrm{DNF}(\tilde{I}^\sigma)\right] &= \sum_{i=1}^{m^2} \mathbb{E}_{s_i \sim \mu_i}\left[\mathrm{DNF}(s_i, \tilde{I}_i^\sigma)\right] \\ &\geq \sum_{i=1}^{m^2} \mathbb{E}_{s \sim \pi^F}\left[\mathrm{DNF}(s, \tilde{I}_i^\sigma)\right] - m^2 \\ &= m^2 \cdot \mathbb{E}_{s \sim \pi^F}\left[\mathrm{DNF}(s, \tilde{I}_1^\sigma)\right] - m^2,\end{aligned}$$

where μ_i denotes the distribution on S after $i \cdot m$ items have been processed.

Now, we bound the difference between the expected values $\mathbb{E}_{s\sim\pi^F}\left[\mathrm{DNF}(s, \tilde{I}_1^\sigma)\right]$ and $\mathbb{E}_{s\sim\pi^F}\left[\mathrm{DNF}(s, I_m^F)\right]$. In order to do this, we use Lemma 93. It is clear that the number of covered bins on such instances is upper bounded by m. It remains to give an upper bound for the total variation distance. One can show that (see e.g. [42])

$$\|\mathcal{L}(\tilde{I}_1^\sigma) - \mathcal{L}(I_m^F)\|_{\mathrm{TV}} \leq \frac{1}{2m}.$$

Therefore, it follows that

$$\mathbb{E}\left[\mathrm{DNF}(\tilde{I}^\sigma)\right] \geq m^2 \cdot \mathbb{E}_{s\sim\pi^F}\left[\mathrm{DNF}(s, I_m^F)\right] - 2m^2.$$

◀

Finally, applying the bound given for (29) in the first part of the proof with $ak + b = n$ and $a = n/B$ yields the result if unit sized items arrive in random order. ◀

5.2.10 Analysis of FF2 in the Random-order Model

▶ **Theorem 83.** *For class-constrained bin covering with unit sized items and parameters k and B it holds* $\mathrm{RR}(\mathrm{FF2}) = 1$.

Proof of Theorem 83. We introduce another notation for this part: For $\ell_1, \ell_2 \in \mathbb{N}_0$ with $\ell_1 \leq \ell_2$ let $[\ell_1 : \ell_2] := \{\ell_1, \ldots, \ell_2\}$.

The proof of this theorem will proceed in three steps:

- At first we will reduce the case of unit sized items arriving in random order to the case where items are drawn independently and identically distributed.
- Then we will model the behavior of the algorithm as a Markov chain, construct another Markov chain as a *comparison chain*, and will relate both chains.
- Finally, we will analyze the growth of the comparison chain.

We remember that FF2 is only designed for the case of unit sized items. Similar to the analysis of Dual-Next-Fit we can define the set of open bin configurations $\mathcal{O}$ for FF2. Let $\mathcal{S} = \{(s, c) \in [0 : B-1] \times \mathcal{P}(\mathbb{N}) : |c| \leq B-1\}$ denote the set of all possible configurations of a single open bin. Due to the design of FF2 we know that if $s = B - i$, where $i \in [k-1]$, we have $|c| = k - i$. Since FF2 is an unbounded-space algorithm, we have

$\mathcal{O} = \bigcup_{i=1}^{\infty} \mathcal{S}^i$. FF2 satisfies the following property: Let $((s_1, c_1), \ldots, (s_u, c_u)) \in \mathcal{O}$, then we have $B - 1 \geq s_1 \geq \ldots \geq s_{u-1} \geq B - (k-1) \geq s_u$ and $c_1 \supseteq c_2 \supseteq \ldots \supseteq c_u$. Finally, for $s \in \mathcal{O}$ and an instance I we denote by $\mathrm{FF2}(s, I)$ the number of covered bins processing I starting with the open bin configuration s.

We observe that FF2 satisfies monotonicity properties:

▶ **Observation 95.**
- *Let I be an instance, and let $\tilde{I}$ be an instance we obtain from I by deleting one item. Then we have* $\mathrm{FF2}(\tilde{I}) \leq \mathrm{FF2}(I)$.
- *Let $s \in \mathcal{O}$ be a configuration of the open bins of* FF2. *Then, we have* $\mathrm{FF2}(s, I) \geq \mathrm{FF2}(I)$.

Now, let $\mathcal{I}$ be the set of items we consider. It follows from the observation that we can assume that the items in $\mathcal{I}$ could be arranged perfectly in bins, that is we have $\mathrm{OPT}(\mathcal{I}) = |\mathcal{I}|/B$.

Let $m = \sup\{u \in \mathbb{N} : u^3 \leq |\mathcal{I}|\}$. Using the same technique as in the proof of Theorem 82 we can reduce the case of items arriving in random order to the case of items that are sampled independently and identically distributed.

▶ **Lemma 96.** *Let F be the discrete perfect-packing distribution induced by $\mathcal{I}$, that is the uniform distribution on $\mathcal{I}$. Then, we have*

$$\mathbb{E}\left[\mathrm{FF2}(I^\sigma)\right] \geq m^2 \mathbb{E}\left[\mathrm{FF2}(I_m^F)\right] - \mathcal{O}(m^2).$$

Let $O(I_m^F)$ denote the number of items in open bins, if FF2 is applied to I_m^F. We can estimate the number of covered bins in terms of the open bins:

$$\mathbb{E}\left[\mathrm{FF2}(I_m^F)\right] \geq \left\lfloor \frac{m - \mathbb{E}\left[O(I_m^F)\right]}{B} \right\rfloor. \tag{33}$$

Then, applying the previous lemma leads us to

$$\begin{aligned} \frac{\mathbb{E}\left[\mathrm{FF2}(I^\sigma)\right]}{\mathrm{OPT}(\mathcal{I})} &= \frac{B \cdot \mathbb{E}\left[\mathrm{FF2}(I^\sigma)\right]}{|\mathcal{I}|} \\ &\geq \frac{B \cdot m^2 \mathbb{E}\left[\mathrm{FF2}(I_m^F)\right] - \mathcal{O}(m^2)}{|\mathcal{I}|} \\ &= \frac{m^3 - m^2 \cdot \mathbb{E}\left[O(I_m^F)\right] - \mathcal{O}(m^2)}{|\mathcal{I}|}. \end{aligned}$$

Therefore, if we can show that $\mathbb{E}\left[O(I_m^F)\right] \in \mathcal{O}(m^{0.5+\epsilon})$, the statement follows. This will be the goal of the remaining part of the proof.

If the items FF2 has to pack are drawn independently and identically distributed with respect to a discrete perfect-packing distribution F, we can model the behavior of FF2 as a Markov chain (X_m) with state-space $S = \mathbb{N}_0 \times \mathcal{O}$. The first component of X_m states the number of covered bins and the second component the open-bin configurations after the m-th item has been processed. In (33) we have seen that we can estimate the performance ratio in terms of the number of items in open bins. Let $Y_m = g(X_m)$ be the stochastic process with state space $\mathbb{N}_0^B$ where Y_m^i, $i \in [B]$, denotes the number of bins with bin level i in X_m.

Now, we construct a chain (Z_m), which we will compare to (Y_m), the *comparison chain*. The state space S of (Z_m) will be also $\mathbb{N}_0^B$. At first we give an informal idea of (Z_m). Let F^* be the distribution on $\mathcal{I} = \{(1, 1), \ldots, (1, k)\}$, where $p((1, 1)) = \frac{B-(k-1)}{B}$, and $p((1, j)) = \frac{1}{B}$ for $j \in [2 : k]$. We can think of Z_m^i as the number of bins of bin level i, if we apply FF2 to $I_m^{F^*}$ with the restriction, that the first $B - (k-1)$ items in a bin are counted as items of color 1.

Now we describe (Z_m) in a formal way: For $i \in [B-1]$ let $v_i = (v_i^1, \dots, v_i^B)$ denote the vector in $\mathbb{N}_0^B$ with $v_i^i = -1$, $v_i^{i+1} = 1$ and the remaining entries are zero. Furthermore, let $v_0 = (1, 0, \dots, 0)$.

Assume we are given a state $(s^1, \dots, s^B) = s \in S$. We set

$$M_1^s = \{i \in [B-(k-1) : B-1] \, : \, s^i > 0\}$$
$$M_2^s = \{i \in [B-k] \, : s^i > 0\} \cup \{0\}.$$

Let p denote the transition kernel of (Z_m). If $M_1^s = \emptyset$, we set

$$p(s + v_{\max M_2^s} | s) = 1.$$

Now assume $M_1^s \neq \emptyset$. Let $m_1, \dots, m_u$ denote the indices contained in M_1^s, and without loss of generality we assume that $m_1 > \dots > m_u$. Furthermore, we set $m_0 := B$. Then, for $j \in [u]$, we set

$$p(s + v_j | s) = \frac{m_{j-1} - m_j}{B} \quad \text{and} \quad p(s + v_{\max M_2^s} | s) = \frac{m_u}{B}.$$

Now we show that we can consider (Z_m) as a minorant for (Y_m). In order to do this, we construct a coupling of the two processes. That is, we construct a process $(\bar{Y}_m, \bar{Z}_m)$ such that $\bar{Y}_m$ has the same distribution as Y_m and $\bar{Z}_m$ has the same distribution as Z_m. Furthermore for $j \in [0 : B-1]$, the following inequality will be satisfied:

$$\sum_{i=0}^{j} \bar{Y}_m^{B-i} \cdot (B-i) \geq \sum_{i=0}^{j} \bar{Z}_m^{B-i} \cdot (B-i). \tag{34}$$

It follows immediately that $\bar{Y}_m$ always has at least as many bins covered as $\bar{Z}_m$.

Now we show how we can construct such a coupling. We will do this using induction. It is clear that (34) is satisfied for the first $B-k$ many items. If there is a j with

$$\sum_{i=0}^{j} \bar{Y}_m^{B-i} \cdot (B-i) > \sum_{i=0}^{j} \bar{Z}_m^{B-i} \cdot (B-i)$$

(34) will also be true after packing the $(m+1)$-th item. Therefore, we call j *critical* if there is equality in (34) and $\bar{Z}_m^{B-(j+1)} > 0$. But then we have $\bar{Y}_m^{B-(j+1)} \geq \bar{Z}_m^{B-(j+1)} > 0$, otherwise (34) would be violated for $j+1$. Let $p_{\bar{Z}}(B-j)$ denote the probability that we put the $(m+1)$-th item in a bin with bin level at least $(B-j)$, and $p_{\bar{Y}}(B-j)$ is defined correspondingly.

Due to construction of $\bar{Z}$ we have $p_{\bar{Z}}(B-j) = j/B$ for $j \in [k-1]$ and $p_{\bar{Z}}(B-j) = 1$ for $j \in [k : B]$. Now consider a bin of X_m containing $B-j$ items. If $j \in [k : B]$ we will put the next item in a bin with bin level $B-j$ with probability 1. Therefore, assume that $j \in [k-1]$. Then, there are j colors in this bin missing. Since we are dealing with discrete perfect-packing distributions, the probability to draw a missing color is at least j/B. Therefore, we see that for all critical j we have

$$p_{\bar{Y}}(B-j) \geq p_{\bar{Z}}(B-j). \tag{35}$$

Let U be a random variable which is uniformly distributed on $[0,1]$. Let $w = \min\{\ell \in [B] \, : \, p_{\bar{Y}}(B-\ell) \geq U\}$ and $w' = \min\{\ell' \in [B] \, : \, p_{\bar{Z}}(B-\ell') \geq U\}$. It follows from (35) that $w \leq w'$. Then, we obtain $\bar{Y}_{m+1}$ from $\bar{Y}_m$ putting the new item into a bin with bin level $B-w$ and we obtain $\bar{Z}_{m+1}$ from $\bar{Z}_m$ putting the new item into a bin with bin level $B-w'$. We see that $(\bar{Y}_{m+1}, \bar{Z}_{m+1})$ also satisfies (34).

Now we will analyze the growth of the comparison chain.

▶ **Lemma 97.** *We have*

$$\mathbb{E}\left[\sum_{i=1}^{B-1} Z_m^i\right] \in \mathcal{O}(m^{0.5+\epsilon}).$$

To prove the statement we use the toolbox of the mathematical machinery for stochastic processes. We will use a statement from [65], which deals with quantifying the growth of trajectories of stochastic processes. We will repeat the statement to make the proof of the lemma self-contained.

Let a be a function, which satisfies the following condition: Suppose that for some $n_a \in \mathbb{Z}_+$ the function $a : [n_a, \infty) \to (0, \infty)$ satisfies the following conditions

1. $x \mapsto a(x)$ is increasing on $x \geq n_a$;
2. $\lim_{x\to\infty} a(x) = \infty$;
3. $\sum_{n \geq n_a} a(n)^{-1} < \infty$.

▶ **Lemma 98** (Theorem 2.8.1 in [65]). *Let Z_m be an $\mathcal{F}_m$-adapted process on $\mathbb{R}_+$. Suppose that there exists a non-decreasing function $f : \mathbb{R}_+ \to \mathbb{R}_+$ and a constant $b \in \mathbb{R}_+$ for which $f(Y_0)$ is integrable, and*

$$\mathbb{E}\left[f(Z_{m+1}) - f(Z_m) \middle| \mathcal{F}_m\right] \leq b \quad \textit{almost surely.,}$$

for all $m \geq 0$. Define the non-decreasing function f^{-1} for $x > 0$ by

$$f^{-1}(x) := \sup\{y \geq 0 \,:\, f(y) < x\}.$$

Furthermore, let a be as described above. Then, almost surely for all but finitely many $m \geq 0$,

$$\max_{0 \leq j \leq m} Z_j \leq f^{-1}(a(2m)).$$

Now we want to apply this statement to prove Lemma 97.

Proof of Lemma 97. At first, we observe that due to the construction of Z_m we have $\sum_{i=1}^{B-k} Z_m^i \leq 1$. Now we show that also for $i \in [B-(k-1) : B-1]$ it is true that $\mathbb{E}\left[Z_m^i\right] \in \mathcal{O}(m^{0.5+\epsilon})$. We want to apply the previous lemma and set $f(x) = x^2$ and $a(x) = x^{1+2\epsilon}$.

An easy computation shows that $\mathbb{E}\left[(Z_{m+1}^i)^2 - (Z_m^i)^2 \middle| Z_m\right] \leq 1$. Then, we see that

$$\max_{0 \leq j \leq m} Y_j^i \leq f^{-1}(a(2m)) \leq 2m^{0.5+\epsilon}$$

almost surely for all but finitely many m. Therefore, we have $\mathbb{E}\left[\sum_{i=1}^{B-1} Z_m^i\right] \in \mathcal{O}(m^{0.5+\epsilon})$. ◀

So the proof of Theorem 83 follows. ◀

5.2.11 An Online Algorithm for Class-constrained Bin Covering with General Item Sizes

▶ **Corollary 84.** *There exists an online algorithm with random-order ratio 1/3 for class-constrained bin covering with general item sizes.*

Proof of Corollary 84. There must be two conditions satisfied to cover a bin:

- The total size of items in a bin is at least 1. We call this the *size-condition.*
- There must be items of at least k different colors. This is the *color-condition.*

We treat the size-condition and the color-condition independently. The idea of the algorithm is rather simple: We treat every third item as a *color-item* and the remaining ones as *size-items*. Let I_s^σ denote the chosen size-items and I_c^σ the color-items. Then we pack the items from I_s^σ using Dual-Next-Fit for classical bin covering and the items from I_c^σ (independent of their size) using FF2. The number of covered bins is then given by $\min\{\mathrm{DNF}(I_s^\sigma), \mathrm{FF2}(I_c^\sigma)\}$.

The proof that this procedure yields a 1/3-competitive algorithm in the random-order model is based on the result for FF2 if items arrive in random order and the use of concentration inequalities. Let $S(I_s^\sigma)$ denote the total size of items in I_s^σ and $\mathrm{OPT}_c(I_c^\sigma)$ denotes the maximum number of covered bins for class constrained bin covering if we assume that the items in I_c^σ are unit sized items and $B = k$.

▶ **Lemma 99.** *We have*

$$\mathbb{P}\left[S(I_s^\sigma) \le \frac{2}{3}\,\mathrm{OPT}(\mathcal{I}) - \sqrt{\mathrm{OPT}(\mathcal{I}) \cdot \log(\mathrm{OPT}(\mathcal{I}))}\right] \le \mathrm{OPT}(\mathcal{I})^{-3/2}, \quad \textit{and} \tag{36}$$

$$\mathbb{P}\left[\mathrm{OPT}_c(I_c^\sigma) \le \frac{1}{3}\,\mathrm{OPT}(\mathcal{I}) - \sqrt{\mathrm{OPT}(\mathcal{I}) \cdot \log(\mathrm{OPT}(\mathcal{I}))}\right] \le k \cdot \mathrm{OPT}(\mathcal{I})^{-3/2}. \tag{37}$$

Proof. We start by showing (36). Without loss of generality we assume that $|\mathcal{I}|$ is a multiple of 3. Otherwise we add items of size 0. Let X_i denote the size of the i-th drawn item in I_s^σ. We have $\mathbb{E}[X_i] = \mathrm{OPT}(\mathcal{I})/|\mathcal{I}|$. Therefore, it follows that $\mathbb{E}[S(I_s^\sigma)] = 2\,\mathrm{OPT}(\mathcal{I})/3$. The random variables X_i are not independent, but we can see them as a sample drawn with respect to sampling without replacement from a finite population. Applying Proposition 7 yields the result.

We assume that $|\mathcal{I}| = \mathrm{OPT}(I)k + b$. To show the second bound we observe that the worst case is given if we assume that there are $\mathrm{OPT}(\mathcal{I})$ items of each of the colors $\{1, \ldots, k\}$ and ignore the colors of the remaining items. Let Y_i be equal to one if the i-th drawn item in I_c^σ is of color 1 and zero otherwise. Then we have $\mathbb{E}[Y_i] = \mathrm{OPT}(\mathcal{I})/|\mathcal{I}|$ and hence, $\mathbb{E}\left[\sum_{i=1}^{|I|/3} Y_i\right] = \mathrm{OPT}(\mathcal{I})/3$. Again applying Proposition 7 this yields that there are at least $\mathrm{OPT}(\mathcal{I}) - \sqrt{\mathrm{OPT}(\mathcal{I}) \cdot \log(\mathrm{OPT}(\mathcal{I}))}$ items of color one in I_c^σ with a probability of at least $\mathrm{OPT}(\mathcal{I})^{-3/2}$. We obtain the result by applying a union to each of the k colors. ◀

Using the previous lemma, the fact that Dual-Next-Fit for classical bin covering is 1/2-competitive and Theorem 83 we conclude that the proposed algorithm has a random-order ratio of 1/3. ◀

6 Conclusions and Open Problems

In this thesis we discussed the performance of online algorithms for variants of bin packing and bin covering with respect to probabilistic performance measures.

In Section 3 we studied the fundamental complexity of classical, cardinality-constrained, and class-constrained bin packing: That is, we dealt with the question if it is possible to design algorithms which are optimal with respect to the performance measures under consideration or if there exist non-trivial lower bounds. This was motivated especially by the following two points:

- The race for the lower bound for the competitive ratio for online classical bin packing [9, 10, 14, 61, 82, 84];
- The existence of an optimal randomized algorithm for the average performance ratio with respect to the set of *all* distributions, which was shown by Rhee and Talagrand 1993 [72].

We took up the approach of Rhee and Talagrand and showed that for cardinality-constrained and class-constrained bin packing there also exists an optimal randomized algorithm for the average performance ratio with respect to the set of all distributions.

Using the same technique it was possible to obtain an optimal algorithm for cardinality-constrained bin packing with respect to the random-order ratio. Then, we combined this approach with packing small items using a First-Fit procedure, which allowed us to obtain for each ϵ greater than zero a randomized algorithm A_ϵ with $\mathrm{RR}(A_\epsilon) \leq 1 + \epsilon$ for classical bin packing. This is a huge step forward to the question of Kenyon for an optimal algorithm with respect to the random-order ratio [54]. Furthermore, this rules out the possibility of a non-trivial lower bound for the random-order ratio. Then, we studied classical bin packing with respect to the partial-permutations model: It turned out that it is necessary that *all* items, and not only a fraction of them, arrive in random order to obtain an optimal algorithm.

For class-constrained bin packing we established a lower bound of $10/9$ for the random-order ratio. So this variant exhibits different complexities studying it with respect to the probabilistic performance measures under consideration. This justifies studying bin packing variants in different stochastic settings.

In this context a number of questions still remain to be answered:

- Is there an online algorithm which is optimal with respect to the random-order ratio for classical bin packing and what does it look like? The replacement procedure by Rhee and Talagrand and our First-Fit-approach fail on packing small items optimally. So there is a new approach necessary to handle this case.
- What is the complexity of the studied bin covering variants with respect to the performance measures under consideration? This question was considered in [26] in the special case that the demand is given by a parameter $B \in \mathbb{N}$, all item sizes are in $\{1, \ldots, B\}$, and the items are drawn independently and identically distributed.
  In the case of general item sizes it is unclear how to transpose the replacement procedure established by Rhee and Talagrand to the case where we have to *cover* bins. Perhaps, it is more promising to try to adapt the approach of Gupta and Radovanovic [46], which is based on gradient descent, to this more general setting.
  Finally, the case of class-constrained bin covering with general item sizes is even more intricate: Is it possible to overcome the difficulty that items have different colors or do non-trivial lower bounds for the probabilistic performance measures exist?
- Eventually, it is interesting to find more bin packing/bin covering variants, which have different complexities in the considered probabilistic settings. A possible candidate could be *dynamic* bin packing, a bin packing variant where items also depart after some time.

In Section 4 we showed that in the bounded-space setting it is possible to beat important worst-case bounds when we consider the random-order ratio:

- There exists a 4-bounded-space algorithm for classical bin packing with a random-order ratio, which is smaller than $h_\infty \approx 1.691$. Here, h_∞ is the lower bound for the competitive ratio of bounded-space algorithms established by Lee and Lee [58], which holds true for *every* bounded-space algorithm – no matter how many bins it is allowed to use.
- For classical bin covering we showed that the random-order ratio of Dual-Next-Fit lies in the interval $[0.502 : 0.\bar{6}]$, proving the conjecture given in [17].

There are huge gaps between the given lower and upper bounds, but this is nothing unusual in this area: For the random-order ratio of Best-Fit in classical bin packing only a lower bound of 1.08 and an upper bound of 1.5 is known [54], and this gap remains open since more than twenty years. Nonetheless, it would be very interesting to tighten these gaps. For Dual-Next-Fit we assume that the upper bound of $2/3$ represents the truth.

Moreover, it is interesting to analyze the performance of bounded-space algorithms with small numbers of bins for classical bin packing in more details: Of special interest is the question of how many bins are necessary to break the worst-case lower bound of h_∞ in the random-order model. It follows from [20] that one bin is not sufficient, but possibly two or three bins are.

Finally, in Section 5 we studied selected heuristics for class-constrained bin packing and bin covering. When we analyzed Next-Fit, First-Fit and the ColorSets-approach for class-constrained bin packing we found that the performance with respect to the average performance ratio and the random-order ratio differs clearly. So the different complexity observed in Section 3 is not only a theoretical phenomenon, but also concerns common heuristics. This underlines for a second time that it is important to consider different probabilistic performance measures.

Then, we turned to the analysis of class-constrained bin covering: Here, we do not observe different behaviors of Next-Fit- and First-Fit-approaches with respect to the average performance ratio and the random-order ratio. Actually, the analyzed heuristics Dual-Next-Fit and FF2 benefit strongly from randomized input (in contrast to their counterparts for class-constrained bin packing). FF2 is even optimal in the case of unit sized items that are revealed in random order. We used this result to show the existence of a randomized offline algorithm with approximation ratio $1/3$ for class-constrained bin covering with general item sizes. As far as we know this is the first algorithm designed for this problem.

Especially, in class-constrained bin covering there are several open questions:

- Even in the case of packing unit sized items online, there is an exponentially large gap between the best known upper bound $\mathcal{O}\left(\log(k)^{-1}\right)$ and the best known algorithm with competitive ratio of $\Omega(1/k)$ [37].
- For the offline version with unit sized items it is known that it is possible to solve it in polynomial time. But what about the complexity of the case of general item sizes? The problem is NP-complete, but do asymptotic (fully) polynomial time approximation schemes exist, and what is possible in the case that the number of different colors is not constant but part of the input?

References

1. Susanne Albers and Michael Mitzenmacher. Average-case analyses of first fit and random fit bin packing. *Random Structures & Algorithms*, 16(3):240–259, 2000.
2. David Arthur, Bodo Manthey, and Heiko Röglin. k-means has polynomial smoothed complexity. In *Proc. of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 405–414, 2009.
3. Eyjolfur Ingi Asgeirsson and Cliff Stein. Bounded-space online bin cover. *Journal of Scheduling*, 12(5):461–474, 2009.
4. Susan F Assmann. *Problems in discrete applied mathematics*. PhD thesis, Massachusetts Institute of Technology, Department of Mathematics, 1983.
5. Susan F Assmann, David S Johnson, Daniel J Kleitman, and JY-T Leung. On a dual version of the one-dimensional bin packing problem. *Journal of Algorithms*, 5(4):502–525, 1984.
6. Luitpold Babel, Bo Chen, Hans Kellerer, and Vladimir Kotov. Algorithms for on-line bin-packing problems with cardinality constraints. *Discrete Applied Mathematics*, 143(1-3):238–251, 2004.
7. János Balogh, József Békési, György Dósa, Leah Epstein, and Asaf Levin. Online Bin Packing with Cardinality Constraints Resolved. In *Proc. of the 25th Annual European Symposium on Algorithms (ESA)*, pages 10:1–10:14, 2017.
8. János Balogh, József Békési, György Dósa, Leah Epstein, and Asaf Levin. A New and Improved Algorithm for Online Bin Packing. In *Proc. of the 26th Annual European Symposium on Algorithms (ESA)*, pages 5:1–5:14, 2018.
9. János Balogh, József Békési, György Dósa, Leah Epstein, and Asaf Levin. A new lower bound for classic online bin packing. *arXiv preprint arXiv:1807.05554*, 2018.
10. János Balogh, József Békési, and Gábor Galambos. New lower bounds for certain classes of bin packing algorithms. *Theoretical Computer Science*, 440:1–13, 2012.
11. Cyril Banderier, René Beier, and Kurt Mehlhorn. Smoothed analysis of three combinatorial problems. In *Proc. of the International Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 198–207, 2003.
12. Rémi Bardenet, Odalric-Ambrym Maillard, et al. Concentration inequalities for sampling without replacement. *Bernoulli*, 21(3):1361–1385, 2015.
13. Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.
14. Donna J Brown. A lower bound for on-line one-dimensional bin packing algorithms. *Coordinated Science Laboratory Report no. UILU-ENG 78-2257, R-864, ACT-19*, 1979.
15. Alberto Caprara, Hans Kellerer, and Ulrich Pferschy. Approximation schemes for ordered vector packing problems. *Naval Research Logistics*, 50(1):58–69, 2003.
16. Barun Chandra. Does randomization help in on-line bin packing? *Information Processing Letters*, 43(1):15–19, 1992.
17. Marie G Christ, Lene M Favrholdt, and Kim S Larsen. Online bin covering: Expectations vs. guarantees. *Theoretical Computer Science*, 556:71–84, 2014.
18. Edward G Coffman Jr, Costas Courcoubetis, Michael R Garey, David S Johnson, Lyle A McGeoch, Peter W Shor, Richard R Weber, and Mihalis Yannakakis. Fundamental discrepancies between average-case analyses under discrete and continuous distributions: A bin packing case study. In *Proc. of the 23rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 230–240, 1991.
19. Edward G Coffman Jr, János Csirik, Gábor Galambos, Silvano Martello, and Daniele Vigo. Bin packing approximation algorithms: survey and classification. *Handbook of combinatorial optimization*, pages 455–531, 2013.

20 Edward G Coffman Jr, János Csirik, Lajos Rónyai, and Ambrus Zsbán. Random-order bin packing. *Discrete Applied Mathematics*, 156(14):2810–2816, 2008.

21 Edward G Coffman Jr, David S Johnson, Peter W Shor, and Richard R Weber. Markov chains, computer proofs, and average-case analysis of best fit bin packing. In *Proc. of the 25th Annual ACM Symposium on Theory of Computing (STOC)*, pages 412–421, 1993.

22 Edward G Coffman Jr, David S Johnson, Peter W Shor, and Richard R Weber. Bin packing with discrete item sizes, part ii: Tight bounds on first fit. *Random Structures & Algorithms*, 10(1-2):69–101, 1997.

23 Edward G Coffman Jr, Kimming So, Micha Hofri, and AC Yao. A stochastic model of bin-packing. *Information and Control*, 44(2):105–115, 1980.

24 János Csirik, Hans Frenk, Gábor Galambos, and Alexander Rinnooy Kan. Probabilistic analysis of algorithms for dual bin packing problems. *Journal of Algorithms*, 12(2):189–203, 1991.

25 János Csirik and David S Johnson. Bounded space on-line bin packing: Best is better than first. *Algorithmica*, 31(2):115–138, 2001.

26 János Csirik, David S Johnson, and Claire Kenyon. Better approximation algorithms for bin covering. In *Proc. of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 557–566, 2001.

27 Janos Csirik, David S Johnson, Claire Kenyon, James B Orlin, Peter W Shor, and Richard R Weber. On the sum-of-squares algorithm for bin packing. *Journal of the ACM*, 53(1):1–65, 2006.

28 János Csirik and Vilmos Totik. Online algorithms for a dual version of bin packing. *Discrete Applied Mathematics*, 21(2):163–167, 1988.

29 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.

30 György Dósa and Leah Epstein. Online bin packing with cardinality constraints revisited. *arXiv preprint arXiv:1404.1056*, 2014.

31 György Dósa and Jiří Sgall. First fit bin packing: A tight analysis. In *Proc. of the 30th International Symposium on Theoretical Aspects of Computer Science (STACS 2013)*, pages 538–549, 2013.

32 György Dósa and Jiří Sgall. Optimal analysis of best fit bin packing. In *Proc. of the International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 429–441, 2014.

33 Aryeh Dvoretzky, Jack Kiefer, Jacob Wolfowitz, et al. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, 27(3):642–669, 1956.

34 Evgenii Borisovich Dynkin. The optimum choice of the instant for stopping a markov process. *Soviet Mathematics*, 4:627–629, 1963.

35 Matthias Englert, Heiko Röglin, and Berthold Vöcking. Worst case and probabilistic analysis of the 2-opt algorithm for the TSP. In *Proc. of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1295–1304, 2007.

36 Leah Epstein. Online bin packing with cardinality constraints. *SIAM Journal on Discrete Mathematics*, 20(4):1015–1030, 2006.

37 Leah Epstein, Csanád Imreh, and Asaf Levin. Class constrained bin covering. *Theory of Computing Systems*, 46(2):246–260, 2010.

38 Leah Epstein, Csanád Imreh, and Asaf Levin. Class constrained bin packing revisited. *Theoretical Computer Science*, 411(34-36):3073–3089, 2010.

39 Leah Epstein and Asaf Levin. Afptas results for common variants of bin packing: A new method for handling the small items. *SIAM Journal on Optimization*, 20(6):3121–3145, 2010.

40 Michael Fekete. Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Mathematische Zeitschrift*, 17(1):228–249, 1923.

41 Carsten Fischer and Heiko Röglin. Probabilistic analysis of the dual next-fit algorithm for bin covering. *arXiv preprint arXiv:1512.04719*, 2015.

42 Carsten Fischer and Heiko Röglin. Probabilistic analysis of the dual next-fit algorithm for bin covering. In *Proc. of the Latin American Symposium on Theoretical Informatics (LATIN)*, pages 469–482, 2016.

43 Carsten Fischer and Heiko Röglin. Probabilistic analysis of online (class-constrained) bin packing and bin covering. In *Proc. of the Latin American Symposium on Theoretical Informatics (LATIN)*, pages 461–474, 2018.

44 Michael R Garey and David S Johnson. *Computers and intractability.* W H Freeman & Co, 2002.

45 Leana Golubchik, Sanjeev Khanna, Samir Khuller, Ramakrishna Thurimella, and An Zhu. Approximation algorithms for data placement on parallel disks. *ACM Transactions on Algorithms (TALG)*, 5(4):34, 2009.

46 Varun Gupta and Ana Radovanovic. Online stochastic bin packing. *arXiv preprint arXiv:1211.2687*, 2012.

47 Rebecca Hoberg and Thomas Rothvoss. A logarithmic additive integrality gap for bin packing. In *Proc. of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2616–2625, 2017.

48 Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

49 Klaus Jansen and Roberto Solis-Oba. An asymptotic fully polynomial time approximation scheme for bin covering. *Theoretical Computer Science*, 306(1-3):543–551, 2003.

50 David S Johnson. *Near-optimal bin packing algorithms.* PhD thesis, Massachusetts Institute of Technology, 1973.

51 David S Johnson. Fast algorithms for bin packing. *Journal of Computer and System Sciences*, 8(3):272–314, 1974.

52 David S Johnson, Alan J Demers, Jeffrey D Ullman, Michael R Garey, and Ronald L Graham. Worst-case performance bounds for simple one-dimensional packing algorithms. *SIAM Journal on Computing*, 3(4):299–325, 1974.

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

54 Claire Kenyon. Best-fit bin-packing with random order. In *Proc. of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 359–364, 1996.

55 Claire Kenyon, Yuval Rabani, and Alistair Sinclair. Biased random walks, lyapunov functions, and stochastic analysis of best fit bin packing. *Journal of Algorithms*, 27(2):218–235, 1998.

56 Kenneth L Krause, Vincent Y Shen, and Herbert D Schwetman. Analysis of several task-scheduling algorithms for a model of multiprogramming computer systems. *Journal of the ACM*, 22(4):522–550, 1975.

57 CC Lee and DT Lee. Robust online bin packing algorithms. Technical report, Technical report 83–03-FC-02, Department of Electrical Engineering and Computer Science, 1987.

58 Chung-Chieh Lee and Der-Tsai Lee. A simple on-line bin-packing algorithm. *Journal of the ACM*, 32(3):562–572, 1985.

59 Tom Leighton and Peter Shor. Tight bounds for minimax grid matching with applications to the average case analysis of algorithms. *Combinatorica*, 9(2):161–187, 1989.

60 David A Levin, Yuval Peres, and Elizabeth L Wilmer. *Markov Chains and Mixing Times.* AMS, 2009.

61 Frank M Liang. A lower bound for on-line bin packing. *Information Processing Letters*, 10(2):76–79, 1980.

62 Denis V Lindley. Dynamic programming and decision theory. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 10(1):39–51, 1961.

63 Bodo Manthey and Rüdiger Reischuk. Smoothed analysis of binary search trees. *Theoretical Computer Science*, 378(3):292–315, 2007.

64 Pascal Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The Annals of Probability*, pages 1269–1283, 1990.

65 Mikhail Menshikov, Serguei Popov, and Andrew Wade. *Non-homogeneous Random Walks: Lyapunov Function Methods for Near-Critical Stochastic Systems.* Cambridge Tracts in Mathematics. Cambridge University Press, 2016.

66 Stephen J Montgomery-Smith. Comparison of sums of independent identically distributed random variables. *arXiv preprint math/9310214*, 1993.

67 Christos H Papadimitriou. *Computational complexity.* John Wiley and Sons Ltd., 2003.

68 Alexander Pruss. A maximal inequality for partial sums of finite exchangeable sequences of random variables. *Proceedings of the American Mathematical Society*, 126(6):1811–1819, 1998.

69 Prakash Ramanan. Average-case analysis of the smart next fit algorithm. *Information Processing Letters*, 31(5):221–225, 1989.

70 Prakash Ramanan, Donna J Brown, Chung-Chieh Lee, and Der-Tsai Lee. On-line bin packing in linear time. *Journal of Algorithms*, 10(3):305–326, 1989.

71 Wansoo T Rhee and Michel Talagrand. Exact bounds for the stochastic upward matching problem. *Transactions of the American Mathematical Society*, 307(1):109–125, 1988.

72 Wansoo T Rhee and Michel Talagrand. On-line bin packing of items of random sizes, ii. *SIAM Journal on Computing*, 22(6):1251–1256, 1993.

73 Thomas Rothvoß. Approximating bin packing within $O(\log(\mathrm{OPT})\log(\log(\mathrm{OPT})))$ bins. In *Proc. of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 20–29, 2013.

74 Steven S Seiden. On the online bin packing problem. *Journal of the ACM*, 49(5):640–671, 2002.

75 Richard Serfozo. *Basics of applied stochastic processes.* Springer Science & Business Media, 2009.

76 Hadas Shachnai and Tami Tamir. Polynomial time approximation schemes for class-constrained packing problems. *Journal of Scheduling*, 4(6):313–338, 2001.

77 Hadas Shachnai and Tami Tamir. Tight bounds for online class-constrained packing. *Theoretical Computer Science*, 321(1):103–123, 2004.

78 Peter W Shor. The average-case analysis of some on-line algorithms for bin packing. *Combinatorica*, 6(2):179–200, 1986.

79 David Simchi-Levi. New worst-case results for the bin-packing problem. *Naval Research Logistics*, 41(4):579–585, 1994.

80 Daniel A Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.

81 Jeffrey D Ullman. *The Performance of a Memory Allocation Algorithm.* Technical report 100 (Princeton University. Dept. of Electrical Engineering. Computer Sciences Laboratory). Princeton University, 1971.

82 André van Vliet. An improved lower bound for on-line bin packing algorithms. *Information Processing Letters*, 43(5):277–284, 1992.

83 Eduardo C Xavier and Flávio K Miyazawa. The class constrained bin packing problem with applications to video-on-demand. *Theoretical Computer Science*, 393(1-3):240–259, 2008.

84 Andrew Chi-Chih Yao. New algorithms for bin packing. *Journal of the ACM*, 27(2):207–227, 1980.

85 Andrew Chi-Chin Yao. Probabilistic computations: Toward a unified measure of complexity. In *Proc. of the 18th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 222–227, 1977.