Interior-point Based Online Stochastic Bin Packing

Varun Gupta
Booth School of Business, University of Chicago, Chicago, IL 60637, varun.gupta@chicagobooth.edu

Ana Radovanović
Google Research, 1245 Charleston Rd., Mountain View, CA 94043, anaradovanovic@google.com

Bin packing is an algorithmic problem that arises in diverse applications such as remnant inventory systems, shipping logistics, and appointment scheduling. In its simplest variant, a sequence of T items (e.g., orders for raw material, packages for delivery) is revealed one at a time, and each item must be packed on arrival in an available bin (e.g., remnant pieces of raw material in inventory, shipping containers). The sizes of items are *i.i.d.* samples from an unknown distributions, but the sizes are known when the items arrive. The goal is to minimize the number of non-empty bins (equivalently waste, defined to be the total unused space in non-empty bins). This problem has been extensively studied in the Operations Research and Theoretical Computer Science communities, yet all existing heuristics either rely on learning the distribution or exhibit $o(T)$ additive suboptimality compared to the optimal offline algorithm only for certain classes of distributions (those with sublinear optimal expected waste). In this paper, we propose a family of algorithms which are the first truly distribution-oblivious algorithms for stochastic bin packing, and achieve $\mathcal{O}(\sqrt{T})$ additive suboptimality for all item size distributions. Our algorithms are inspired by approximate interior-point algorithms for convex optimization. In addition to regret guarantees for *i.i.d.* sequences, we also prove a family of novel regret bounds for general *non-i.i.d.* input sequences, including guarantees for locally adversarially perturbed i.i.d. sequences. To the best of our knowledge these are the first such results for non-i.i.d. and non-random-permutation input sequences for online stochastic packing.

Key words: Bin packing, Primal-Dual algorithm, penalized Lagrangian, semi-adversarial input

1. Introduction

Bin packing is one of the oldest resource allocation problems and has received considerable attention due to its practical relevance. In the classical offline version of static bin packing, a list of scalar item sizes $\{s_1, s_2, \ldots, s_T\}$ has to be partitioned into the fewest number of partitions each summing to at most B (the bin size). In the still more challenging online version, the list of item sizes is revealed one at a time, and the items must be irrevocably assigned to a bin on arrival. Online bin packing appears as a motif in many operations research problems, of which we give a small sampling below:

1. Remnant Scheduling/Cutting Stock: In Adelman and Nemhauser (1999), the authors cite the example of a fiber optic cable manufacturer which produces cables of fixed lengths. Orders for customer-specified lengths arrive and must be served online from the available inventory. The goal is to serve the demand while minimizing the rate of production of cables, or equivalently the scrap rate of remnant inventory. In this context, bins correspond to the fixed length cables, and items correspond to the customer orders which must be packed (i.e., cut) online and irrevocably.

2. Appointment scheduling: Requests for appointments arrive online and must be scheduled in a future day with available slots. Here appointments correspond to items, and the office hours during a working day correspond to bins. In addition to the remnant scheduling example, one novel feature in appointment scheduling problem is that bins depart periodically, and the goal is to minimize the time until appointment. However, at its core is a bin packing problem.

3. Transportation logistics: Items to be shipped arrive and are queued at a transshipment center. Bins correspond to shipping containers (which arrive either exogeneously, or endogeneously on demand), and must be packed using queued items. The goal is to minimize some combination of shipping costs (the number of bins used) and holding costs (number of items waiting to be packed).

The common thread in all the three examples above is that items are packed irrevocably, and do not depart from the bin once packed (the bin may depart with the items). This is known as the *static bin packing* problem. In this paper we will develop algorithms for the first variant of static online bin packing: an unbounded number of bins are available, items are packed irrevocably on arrival into a feasible bin, and the goal is to minimize the number of bins used.

For adversarially generated instances (that is worst-case analysis), offline bin packing problem is NP-hard, but good approximation algorithms exist for one-dimensional packing (Karmarkar and Karp (1982), Rothvoß (2013)). For online bin packing, simple algorithms such as Best Fit are the state-of-the-art and known to use at most $\frac{17}{10}$ times the optimal number of bins (Johnson et al. (1974)). Driven by practical considerations and the obstacles described above, a common approach has been to make stochastic assumptions on the problem instance – the number of items to be packed, T, is much larger than the number of items that can fit in a bin, and the item sizes are assumed to be an *i.i.d.* sequence from some distribution F. Further, usually the item sizes and bin sizes are assumed to integers. The performance of an online packing heuristic is measured by the the expected difference between the number of bins used by the online heuristic, and the optimal-in-hindsight algorithm (ideally, we desire the suboptimality gap, or regret, that grows as $o(T)$).

Our main driving question in the paper is: *Are there simple distribution-oblivious online packing algorithms that perform near optimally for* i.i.d. *inputs? What robustness guarantees can we prove for such algorithms on non-stationary input sequences?* In addition to the intuitive appeal of heuristics with 'fewer moving parts,' these questions also address the value of learning and information in devising control policies.

There has been an extensive study of heuristics for stochastic online bin packing. One line of research has focused on analysis of common heuristics (e.g., Best Fit (BF)), and identifying item size distributions for which these can be optimal, or provably suboptimal. On the algorithmic front, heuristics have been proposed which are asymptotically optimal as the number of items

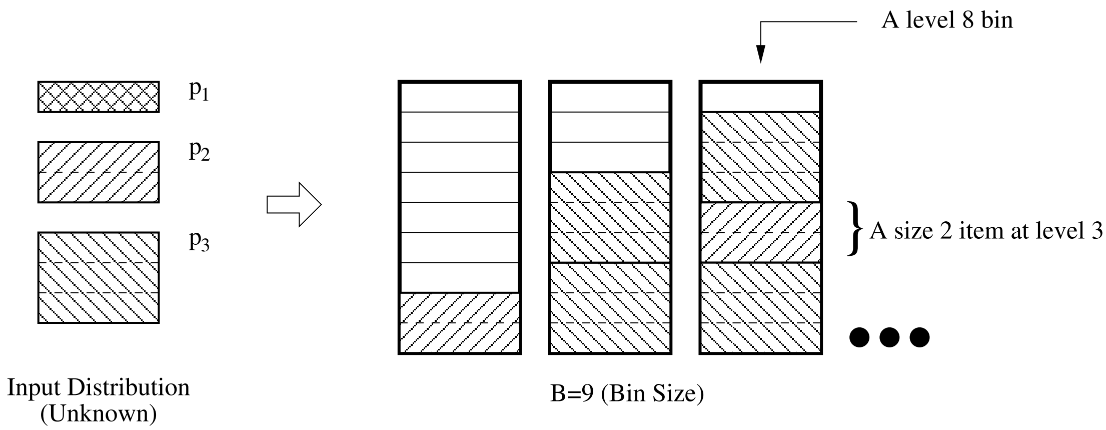


Figure 1 A bin packing instance and nomenclature

grows. However, to the best of our knowledge, all known heuristics explicitly learn the item size distribution and at some level involve solving a linear program (LP) to tune the heuristic. In this paper we present simple algorithms that are distribution-oblivious and asymptotically optimal. Our algorithms are motivated by an approximate Interior-Point (Primal-Dual) solution of the bin packing LP, but surprisingly, has not appeared in the literature. Further, we also prove a novel family of regret guarantees for non-*i.i.d.* input sequences which yield as a corollary regret for sequences which have been locally and adversarially perturbed starting from an *i.i.d.* sequence before being presented to the online packing algorithm.

2. Model Notation and Definitions

A sequence $Y = \{Y_1, Y_2, \ldots, Y_T\}$ of items is packed online using an algorithm A. The sizes Y_t are *i.i.d.* samples from $[J] = \{1, 2, \ldots, J\}$. We will sometimes also refer to the size of an item as its type. We will denote the probability of type/size j items by p_j, abbreviate the item size distribution by $F \triangleq \{p_1, \ldots, p_J\}$, and by U_F the set of item types with strictly positive probability under F (i.e., the support of F). All bins have capacity B – a scalar integer, and we assume $J = B - 1$ without loss of generality. A bin is called *level* h bin if the sizes of the items packed in the bin sum to h. See Figure 1 for a visual illustration of the definitions so far.

For a given packing P of items into bins, $N_h(P)$ will denote the number of level h bins in P. Let $P_F^A(t)$ denote the packing after applying algorithm A to t items generated from distribution F. With abuse of notation, we use $N_h^A(t)$ to denote the number of bins of level h in $P_F^A(t)$.

Metric: A natural metric to minimize is the total number of bins open in the packing P:

$$N(P) \triangleq \sum_{h=1}^{B} N_h(P).$$

There is an alternate but equivalent performance metric that is more commonly used in the one-dimensional level model – the *waste* of packing P:

$$W(P) \triangleq \frac{1}{B}\sum_{h=1}^{B-1}(B-h)\cdot N_h(P).$$

That is, the waste of packing P is the unused space in the bins open in P (measured in units of bins). The expected waste rate of algorithm A on distribution F is given by as:

$$\mathbf{E}\big[W_F^A(T)\big] \triangleq \mathbf{E}\big[W(P_F^A(T))\big]. \tag{1}$$

Similarly, we define the expected number of bins used by algorithm A on distribution F:

$$\mathbf{E}\big[N_F^A(T)\big] \triangleq \mathbf{E}\big[N(P_F^A(T))\big]. \tag{2}$$

We use $\mathbf{E}[W_F^{OPT}(T)]$ and $\mathbf{E}[N_F^{OPT}(T)]$ to denote the expected waste and expected number of bins used in the optimal offline packing.

A classification of item-size distributions: To describe the performance of the current state-of-the-art algorithm for stochastic online bin packing, we will need the following classification result of Courcoubetis and Weber (1986): Any discrete item-size distribution F falls in one of three categories based on the asymptotic growth rate of waste of the optimal offline algorithm ($\mathbf{E}[W_F^{OPT}(T)]$) as a function of T:

1. Linear Waste (LW) : $\mathbf{E}[W_F^{OPT}(T)] = \Theta(T)$, e.g., $B=9, F=\{p_2=0.8, p_3=0.2\}$
2. Perfectly Packable (PP) : $\mathbf{E}[W_F^{OPT}(T)] = \Theta(\sqrt{T})$, e.g. $B=9, F=\{p_2=\frac{3}{4}, p_3=\frac{1}{4}\}$
3. PP with Bounded Waste (BW) : $\mathbf{E}[W_F^{OPT}(T)] = \Theta(1)$, e.g. $B=9, F=\{p_2=0.5, p_3=0.5\}$

The intuition for the above classification is the following: We can represent a feasible packing of a bin as a vector $x \in \mathbb{N}^J$ where the jth components of x is the number of size j items in the bin. For example $x = [0\ 0\ 2]$ represents the middle bin in Figure 1 with two items of size 3 and no other items. Let $\mathcal{X}^*$ represents the set of vectors which are perfectly packed (that is, level B bins). For example, for $B=9$, the vectors $[0\ 0\ 3]$ and $[0\ 3\ 1]$ are in $\mathcal{X}^*$. The set of vectors $\mathcal{X}^*$ generate a convex cone representing item frequency vectors which can be packed (with fractional bins allowed) with zero waste. A distribution F in the interior (or relative interior if the support U_F is not $[J]$) of this cone is a Bounded Waste distribution because an empirical sample of T items from F remains in the interior of this set after discarding $\mathcal{O}(1)$ items. A distribution that is outside the convex cone generated by $\mathcal{X}^*$ will be a Linear Waste distribution. Distributions on the boundary of the convex cone will have $\mathcal{O}(\sqrt{T})$ waste since $\Theta(\sqrt{T})$ items must be discarded from an empirical sample of size T so that the remaining list of items can be packed with zero waste. (In Section 8, we add to this classification by introducing a rather broad class of distributions which we call *Corner Point (CP)* distributions. In fact, CP distributions turn out to be universal in the sense that any random perturbation of a given distribution is a CP distribution.)

3. Prior Work and Our Contributions

The relevant prior literature can be partitioned into algorithms which are distribution-aware or actively learn the distribution, and distribution-oblivious algorithms.

Distribution-aware online packing: Adelman and Nemhauser (1999) consider the problem of minimizing scrap for remnant scheduling (also called the 1-d cutting stock problem) which, as mentioned earlier, is a rephrasing of the bin packing problem. The authors propose an algorithm that learns the item size distribution, and uses the duals of a bin packing Linear Program (LP) while making packing decisions. Rhee and Talagrand (1993) propose a packing heuristic which uses all the item sizes seen so far to form a bin packing LP relaxation and prove that when the item sizes are *i.i.d.* from a general distribution (the support of the distribution can be continuous), their algorithm has regret $\mathcal{O}(T^{1/2}\log^{3/4} T)$. Another relevant work is by Iyengar and Sigman (2004) where the authors devise a control policy for a loss network based on solving an offline LP, and then controlling the system online so as to minimize the deviation from the solution of the LP. However the authors do not explicitly describe the application of their algorithm to static bin packing.

Distribution-oblivious online packing: Most of the work on analysis of distribution-oblivious algorithms for stochastic bin packing has been carried out in the theoretical computer science community, beginning with analysis of First Fit and Best Fit heuristics for which worst-case performance in non-stochastic settings were known from earlier. When the bin size is 1 and item size distribution is Unif$[0,1]$, Shor (1986) proved that the expected waste under First Fit (pack in the oldest feasible bin) grows as $\Theta(T^{\frac{2}{3}})$. For Best Fit (pack in the fullest feasible bin), Leighton and Shor (1986) proved this to be $\Theta(T^{\frac{1}{2}}\log^{\frac{3}{4}} T)$. Finally, Shor (1991) proposed a scheme that achieves the lower bound of $\Theta(T^{\frac{1}{2}}\log^{\frac{1}{2}} T)$. For discrete item sizes, when the item sizes are uniformly distributed over $\left\{\frac{1}{B},\frac{2}{B},\ldots,\frac{J}{B}\right\}$, Coffman et al. (1991) proved the expected waste for $J=B$ or $J=B-1$ grows as $\Theta(TB^{\frac{1}{2}})$ for First Fit, and $\Theta(T^{\frac{1}{2}}\log B)$ for Best Fit. For $J=B-2$, bounded expected waste for Best Fit was proved by Kenyon et al. (1996), and for First Fit (using Random Fit as an intermediate step) by Albers and Mitzenmacher (1998). Kenyon and Mitzenmacher (2000) proved that the waste under Best Fit is linear when $J=\alpha B$, $\frac{99}{150}<\alpha<\frac{100}{150}$, and B large enough, but is conjectured to hold for all $0<\alpha<1$. (Interestingly, Best Fit has linear expected waste even for the benign case of $B=6$ and items of size $2,3$ with equal probability, but this appears to not have been a compelling reason to seek alternatives to BF.)

Sum of Squares (SS) rule Csirik et al. (1999, 2006): The SS heuristic is in some sense the state-of-the-art bin packing policy when item sizes and bin size B are integral. It is almost distribution-oblivious, and nearly universally optimal for all distributions F, and works as follows: Suppose $P(t-1)$ represents the current packing after seeing $t-1$ items. On arrival of the tth item, it is packed in a feasible bin so as to minimize the value of the following potential function of the resulting packing $P(t)$:

$$ss(P(t)) \triangleq \sum_{h=1}^{B-1} N_h(t)^2.$$

Csirik et al. (2006) prove that for PP distributions, the waste under SS is indeed $\mathcal{O}(\sqrt{T})$. Further, for BW distributions, the waste of SS is $\mathcal{O}(\log T)$ which can be reduced to $\mathcal{O}(1)$ by learning the support of the distribution. However, for linear waste distributions SS achieves an $\Theta(T)$ additive suboptimality. That is, SS is not asymptotically optimal for LW distributions. As a result, authors propose to tune the policy by introducing 'phantom' items of size 1 at the correct rate (the smallest rate so that the distribution becomes perfectly packable), but this rate is determined by learning the distribution F and solving an LP. Therefore, SS is not a universally distribution-oblivious packing algorithm. Our proposed heuristics obtains $\Theta(\sqrt{n})$ additive suboptimality for all distributions while being truly 'blind', and therefore is the first asymptotically optimal and universally distribution-oblivious algorithm for stochastic online bin packing.

Online Convex Optimization (OCO) for Online Packing/Covering Problems: A somewhat related thread of research is the literature exploiting online convex optimization tools such as Online Mirror Descent and Multiplicative Update algorithm to solve online packing and covering problems (e.g., Gupta and Molinaro (2014), Agrawal and Devanur (2015)). In online packing and covering problems, an item Y_t is associated with a set of feasible actions $A_t \subset \Re^m$. The algorithm must choose an action $a_t \in A_t$ for Y_t without knowing the sequence of future arrivals, while obeying packing/covering constraints on $\sum_t a_t$ and maximizing a reward function. The arrival models considered in such papers are either *i.i.d.* from an unknown distribution, or random permutation of a possibly adversarially generated sequence of items. The bin packing problem considered in the current paper can be cast in this framework by associating with each item a set of action vectors A_t denoting feasible placements, but there is a crucial difference: in our setting the set of feasible actions is state-dependent. In fact, we believe that robustness results similar to what we prove for non-*i.i.d.* sequences can be carried over to Online Packing/Covering Problems handled in OCO framework.

Static packing models with bin departures: Although not the object of study in the present paper, we briefly mention the literature on static bin packing where bins arrive and/or depart.

Coffman and Stolyar (2001) study a model where at items arrive continuously and queue up. At discrete time instants ($t = 0, 1, 2, \dots$) a bin arrives, is filled using the items currently in queue using some packing scheme (e.g. Best Fit (pack the largest item, then the next largest and so on), First-Fit (try to pack the oldest item)), and then the bin departs immediately. The authors prove sufficient stability conditions for discrete item sizes with symmetric distributions. Gamarnik

(2004) studies the stability for general item size distributions via Lyapunov functions and provide a numerical algorithm for checking stability to arbitrary precision. Gamarnik and Squillante (2005) further investigate the steady-state behavior of Best Fit via Lyapunov function analysis and matrix analytic techniques. In particular, they find that the sufficient condition for stability of symmetric distributions do not carry over to asymmetric distributions. Shah and Tsitsiklis (2008) study the lower and upper bounds for asymptotic order of growth rate of queue length in heavy traffic under symmetric distributions.

Lelarge (2007) studies a model with an infinite collection of bins where items are packed on arrival, and the oldest bin departs at discrete time steps. The performance metric investigated is the number of partially filled bins, and the total size of items packed in the partially filled bins. The main result is that for symmetric item size distributions, both First Fit (FF) and Best Fit (BF) are stable, but the volume of queued items is asymptotically larger under FF in heavy traffic.

3.1. Summary of Contributions and Outline

1. **A new online bin packing algorithm:** In Section 4 we present our **Primal-Dual (PD) family** of online bin packing algorithms. Our algorithms are motivated by using gradient descent to solve an Interior-point relaxation of the bin packing LP, and pack items so as to greedily minimize a penalized-Lagrangian. Choosing different barrier functions (Interior-point view), or penalty functions (penalized Lagrangian view), result in different algorithms in the family. For the configuration model with exponential penalty function, this entails greedily minimizing:

$$\mathcal{L}_{\exp}(\mathbf{N}(\mathbf{t})) = \sum_{h=1}^{B} N_h(t) + \frac{1}{\epsilon(t)} \sum_{h=1}^{T} e^{-\epsilon(t) \cdot N_h(t)}$$

We prove that for the appropriate choice of $\epsilon(t) = \Theta(\frac{1}{\sqrt{t}})$, the algorithm achieves

$$\mathbf{E}\big[N_F^{PD}(T)\big] = \mathbf{E}\big[N_F^{OPT}(T)\big] + \mathcal{O}(\sqrt{T})$$

for all discrete distributions F. (Theorems 1-3)

Our algorithms can equivalently be interpreted as online mirror ascent to solve the dual maximization problem, where the choice of barrier/penalty functions now map to choice of distance generating functions on the space of duals of the bin packing LP. We also provide a nice interpretation of our algorithms as "fixing" the SS algorithm to work for LW distributions.

2. **Bounded inventory guarantees:** The PD algorithm of Theorem 1 keeps all the bins created in its working inventory ($\Theta(T)$), which may not be desirable. We prove that if the online algorithm is only allowed to keep at most I bins open at any time, then a modified PD algorithm yields a packing with $\mathbf{E}[N_F^{PD}(T)] = (1 + \mathcal{O}(1/I))\,\mathbf{E}[N_F^{OPT}(T)]$, and that there are distributions for which no online algorithm, even distribution-aware, can have a competitive ratio $(1 + o(1/I))$. (Theorems 4-5)

3. **Guarantees against non-*i.i.d.* input:** In Section 6 we turn to the setting where item size sequence $\{Y_1, Y_2, \ldots, Y_T\}$ need not be *i.i.d.* from a single distribution F. We will not try to summarize the most general result we can prove (Theorem 6), but a corollary follows: Let $\{Y_1, \ldots, Y_T\}$ be generated by starting from an *i.i.d.* sequence and then being perturbed by an adversary under the constraint that no item is moved more than $W/2$ positions from its initial location in the sequence. Then,
   (a) if W is unknown but $W = o(\sqrt{T})$: $\mathbf{E}[N^{PD}(T)] = \mathbf{E}[N^{OPT}(T)] + \mathcal{O}(T^{3/4}W^{1/2})$;
   (b) if W is known and $W = o(T)$: $\mathbf{E}\left[N^{PD'}(T)\right] = \mathbf{E}[N^{OPT}(T)] + \mathcal{O}(T^{2/3}W^{1/3})$.
   where in PD' the parameter $\epsilon(t)$ is optimized with the knowledge of W. (Corollary 3)
4. **Configuration model for bin packing:** In Section 7 we generalize our algorithms to *configuration model*: here items are assigned types in the set $[J] = \{1, 2, \ldots, J\}$. Unlike in one-dimensional packing where feasible ways to pack a bin were given by implicit capacity constraints, in the configuration model we are explicitly given a set $\mathcal{C}$ of feasible configurations. This setting is flexible enough to model multi-dimensional bin packing, vector packing, and minimum cost packing with heterogeneous bin types.
5. **Corner-Point distributions:** Finally, in Section 8 we extend the Courcoubetis and Weber (1986) classification of distributions by introducing the so-called Corner Point (CP) distributions. CP distributions are universal in the sense that any small random perturbation of a distribution gives a CP distribution with probability 1. We conjecture that if the items $\{Y_1, \ldots, Y_T\}$ are generated from a sequence $\{F_1, F_2, \ldots, F_T\}$ such that each of F_t is a CP distribution "belonging to the same corner point", then setting $\epsilon(t) = \Theta(\log^{1+\beta} t)$ (for arbitrary $\beta > 0$) in the PD algorithm gives $\mathcal{O}(\log^{1+\beta} T)$ regret. That is, even though the distribution may be LW as per Courcoubetis and Weber (1986) criterion, we can pack the sequence online while only using $\mathcal{O}(\log^{1+\beta} T)$ more bins than the optimal-in-hindsight algorithm. We provide intuition and experimental evidence for this conjecture. (Conjecture 1)

4. Online Stochastic Bin Packing with i.i.d. arrivals

In this section we present our algorithm for stochastic online bin packing that achieves additive $\mathcal{O}(\sqrt{T})$ suboptimality for *i.i.d.* item sizes for all distributions F. As mentioned before there are no simple, *distribution-oblivious* bin packing heuristic known that are asymptotically optimal for linear waste and perfectly packable distributions simultaneously. The Sum-of-Squares (SS) heuristic achieves a constant factor larger waste than the optimal for LW distributions. All attempts to improve this performance so far have relied on the knowledge of distribution F (Csirik et al. (2006), Courcoubetis and Weber (1986)) or its learning in the process of packing (Rhee and Talagrand (1993)). Further, there are no heuristics for multi-dimensional stochastic bin packing. Our algorithms fill this gap.

In Section 4.1, we present the Linear Program (LP) for minimum cost fractional packing, and propose a general recipe for converting this LP into an online packing algorithm by transforming it to an interior-point/penalized-Lagrangian problem. In Section 4.2 we present our theorems on performance guarantee of the proposed algorithm. In Section 4.3 we present the result on competitive ratio of our online packing algorithm when there is a bound on the number of bins allowed to be kept open.

4.1. Bin-packing LP and Primal-Dual algorithms

We begin with the Linear Program for the following offline one-dimensional bin packing problem (Csirik et al. (2006)): Given the distribution $F = \{p_1, p_2, \ldots, p_J\}$ and bin size B, what is the average number of bins used per item in the optimal packing?

To solve this problem, we define the following decision variables:

$$v(j,h) = \text{fraction of overall jobs which are of type } j \text{ and are packed at level } h \text{ in their bin.}$$

With the above, we are led to the following LP:

$$\begin{aligned}
& b(F) = \min_{\{v(j,h)\}} \sum_j v(j,0) && (\mathbf{P_{1d-level}}) \\
\text{subject to} \quad & \forall h \in [1, B-1] \; : \; \sum_j v(j, h-j) \geq \sum_j v(j,h) && \textit{(no floating items)} \\
& \forall j \in [J] \; : \; \sum_h v(j,h) = p_j && \textit{(mass balance)} \\
& \forall j \in [J]; h \in [B-j] \; : \; v(j,h) \geq 0
\end{aligned}$$

The expression:

$$n_h \triangleq \sum_j v(j, h-j) - \sum_j v(j,h)$$

denotes the number of level h bins in the optimal fractional packing, and the *(no floating items)* constraint says that $n_h \geq 0$ for all h. The *(mass balance)* constraint says that the total mass of type j items must equal p_j. The objective function counts the fraction of overall items which are packed at level 0, and hence the number of bins $\sum_{h=1}^{B} n_h$. Also define:

$$w(F) = b(F) - \frac{\sum_j j \cdot p_j}{B} \tag{3}$$

which denotes the waste per item under optimal packing.

Our approach is quite straightforward in hindsight: Rather than learn the distribution F and then solve the bin packing LP as has been done in the past, we transform this LP into an interior-point/Lagrangian type objective function by using barrier/penalty functions for the *(no floating* *items)* constraint. We then perform stochastic gradient descent to solve this problem, (the stochasticity coming from the random arrival of items). Our algorithms also have a Primal-Dual interpretation (hence the name PD), but unlike typical Primal-Dual algorithms we do not explicitly maintain dual variables. Instead, the duals are implicitly tracked via a map from the primal variables. We explain this in the following subsection where we begin with a general template for design of Interior point based Primal-Dual algorithms to build intuition, and then present our algorithm and guarantees formally. (See also Section 5.2 for a dual ascent using mirror map interpretation.)

4.1.1. A general template for Primal-Dual algorithms Consider the following convex minimization problem:

$$\text{minimize } f(x)$$
$$\text{subject to } g(x) \leq 0.$$

The interior point/penalty function approach to optimization is to convert the constrained optimization problem into an unconstrained optimization by imposing a strictly convex increasing penalty function $\Phi_\epsilon(\cdot)$ on the constraint and moving this penalty into the objective function:

$$\text{minimize} \quad \mathcal{L}(x) = f(x) + \Phi_\epsilon(g(x))$$

Given the optimal solution x_ϵ^* to the above optimization problem, an approximation to the value of the dual λ_g for the constraint $g(x) \leq 0$ can be obtained by comparing the first order optimality condition for the unconstrained problem:

$$\left[\nabla f + \frac{\partial \Phi_\epsilon(g)}{\partial g} \cdot \nabla g\right]_{x_\epsilon^*} = 0$$

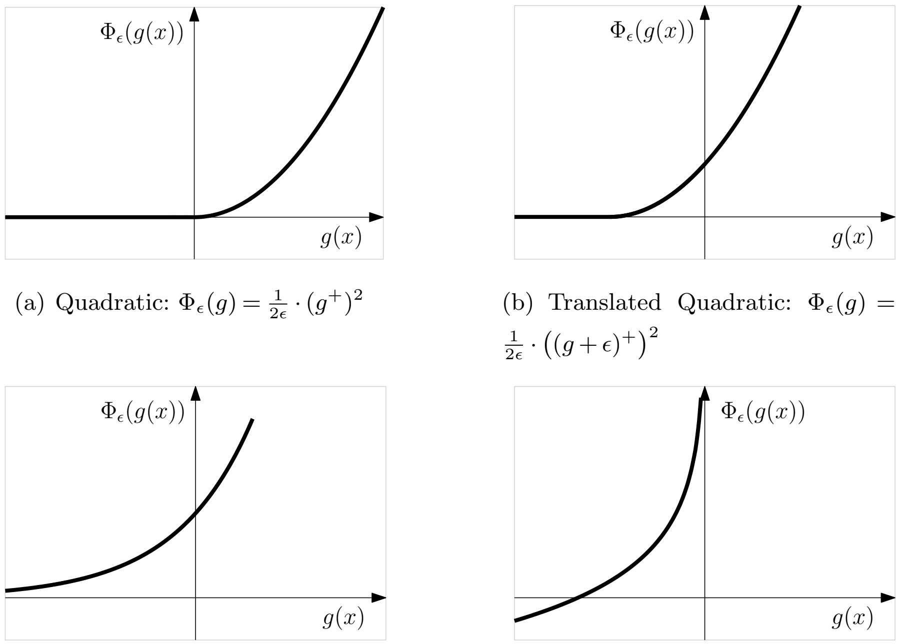
to the KKT stationarity condition for the constrained problem:

$$[\nabla f + \lambda_g \cdot \nabla g]_{x^*} = 0$$

as $\lambda_g \approx \left.\frac{\partial \Phi_\epsilon(g)}{\partial g}\right|_{g(x_\epsilon^*)}$.

Applied to our setting, the function $f(x)$ will correspond to the objective function $\sum_h n_h$ of $\mathbf{P_{1d-level}}$, and $g(x) \leq 0$ will correspond to the constraints $n_h \geq 0$. In typical interior-point algorithms, the unconstrained relaxations are solved using a Newton step starting from the solution of the previous iteration, and ϵ is decreased so that $\Phi_\epsilon(g) \to \mathbf{1}_{\{g \leq 0\}}$. In approximate interior-point iterations, the unconstrained optimization is solved using proximal descent. Our online algorithm is closer to the latter view, except that the direction of descent is constrained by the item type seen by the algorithm as well as the current packing, which results in stochastic gradients.

Now depending on our choice of penalty function $\Phi(\cdot)$, there is a full menu of relaxations, each with its own mapping of primal to dual variables (see Figure 2):



(a) Quadratic: $\Phi_\epsilon(g) = \frac{1}{2\epsilon} \cdot (g^+)^2$

(b) Translated Quadratic: $\Phi_\epsilon(g) = \frac{1}{2\epsilon} \cdot \left((g+\epsilon)^+\right)^2$

(c) Exponential: $\Phi_\epsilon(g) = \epsilon \cdot e^{g/\epsilon}$

(d) log-barrier: $\Phi_\epsilon(g) = -\epsilon \cdot \log(-g)$

Figure 2 Illustration of four barrier/penalty functions for the constraint $g(x) \leq 0$.

- Quadratic:

$$\mathcal{L}_{\text{quad}}(x) = f(x) + \frac{1}{2\epsilon}\left(g(x)^+\right)^2 \qquad\qquad \mathit{Dual}: \lambda_g = \frac{g(x)^+}{\epsilon}$$

For the dual variables to drive the algorithm, we *must violate the primal constraints*. That is, we always have a primal infeasible solution. Primal-dual heuristics with quadratic penalty are common in controlling queueing systems (e.g., Tassiulas and Ephremides (1992), Stolyar (2005)) because queues essentially are temporary violations of capacity constraints and map to the corresponding duals.

- Translated Quadratic:

$$\mathcal{L}_{\text{tquad}}(x) = f(x) + \frac{1}{2\epsilon}\left((g(x)+\eta)^+\right)^2 \qquad\qquad \mathit{Dual}: \lambda_g = \frac{(g(x)+\eta)^+}{\epsilon}$$

A small variation over Quadratic penalty which allows duals to be non-zero even when the constraint is satisfied, but yields zero duals when constraints are η-far from violation.
In Section 5.1, we show that the Primal-Dual algorithms corresponding to Quadratic and Translated Quadratic penalty functions can be interpreted as two different 'fixes' to the Sum-of-Squares heuristic.

- Exponential penalty:

$$\mathcal{L}_{\exp}(x) = f(x) + \epsilon \cdot e^{\frac{g(x)}{\epsilon}} \qquad\qquad \mathit{Dual}: \lambda_g = e^{\frac{g(x)}{\epsilon}}$$

The dual variables are always non-zero even when the primal solution is feasible. Exponential duals are very popular for worst-case (non-stochastic) online packing and covering problems (e.g., Awerbuch and Khandekar (2008), Plotkin et al. (1995), Buchbinder et al. (2007)), and in prediction with experts' advice. Our main PD heuristic is precisely the Lagrangian relaxation of $\mathbf{P_{1d-level}}$ with exponential penalty. This method is also equivalent to Online Mirror Descent algorithm for maximizing the dual of $\mathbf{P_{1d-level}}$ with entropy regularizer as we explain in Section 5.2.

- log-barrier:

$$\mathcal{L}_{\log}(x) = f(x) - \epsilon \cdot \log(-g(x)) \qquad\qquad Dual: \lambda_g = -\frac{\epsilon}{g(x)}$$

The solution is constrained to be always primal feasible. Since, log is a self-concordant barrier function, log-barriers offer provable convergence guarantees for Newton-Raphson iteration. Therefore they are often used in interior point algorithms for convex optimization (see Boyd and Vandenberghe (2004)). We will not discuss log-barrier based Primal-Dual heuristics in this paper as they give worse regret guarantees.

In each case, as $\epsilon, \eta \to 0$, the penalty function approaches the barrier penalty, and ϵ, η control the violation of constraints (for quadratic penalty), or the loss in objective function (for exponential, translated quadratic, and log-barrier).

4.1.2. The Primal-Dual algorithm for bin packing In this section we focus on the Exponential penalty function based interior point relaxation, and formally develop the corresponding Primal-Dual packing algorithm. The interior point relaxation with Exponential penalty function for $\mathbf{P_{1d-level}}$ is given by:

$$\mathcal{L}_{exp}(n) = \sum_{h=1}^{B} n_h + \kappa\epsilon \sum_{h=1}^{B-1} e^{-n_h/\epsilon}$$

Recall that we denote the number of bins of level h at time t by $N_h(t)$, whereas $n_h \sim \frac{N_h(t)}{t}$. Substituting and multiplying throughout by t:

$$\mathcal{L}_{exp}(N(t)) = \sum_{h=1}^{B} N_h(t) + \kappa\epsilon t \sum_{h=1}^{B-1} e^{-\frac{N_h(t)}{\epsilon t}}$$

which we write more generally as:

$$\mathcal{L}_{exp}(N(t)) = \sum_{h=1}^{B} N_h(t) + \frac{\kappa}{\epsilon(t)} \sum_{h=1}^{B-1} e^{-\epsilon(t) N_h(t)}$$

The proposed Primal-Dual algorithm places arriving items so as to greedily minimize the above penalized-Lagrangian. Noting that $\frac{\partial \mathcal{L}(t)}{\partial N_h(t)} = 1 - \kappa e^{-\epsilon(t) N_h(t)}$ for $1 \le h \le B-1$, and 1 for $h = B$ gives Algorithm 1. We discuss the settings for κ and $\epsilon(t)$ in Theorems 1-2.

Algorithm 1 PD-exp for one-dimensional level model

for $t = 1, 2, \ldots$ **do**

Observe item type Y_t (say, $Y_t = j$)

Define level potentials:

$$V_h(t-1) := 1 - \kappa e^{-\epsilon(t) N_h(t-1)} \quad (h = 1, \ldots, B-1)$$
$$V_0 := 0 \ ; \quad V_B := 1$$

Place arriving item in a level h^* bin to create a level $(h^* + j)$ bin where:

$$h^* = \operatorname*{arg\,min}_{h: N_h(t-1) > 0} V_{h+j}(t-1) - V_h(t-1)$$

end for

Simulation experiments: Figure 3 shows a comparison of SS and PD-exp Algorithm 1 for three distributions. [1] To highlight $\mathcal{O}(\sqrt{T})$ regret of PD-exp, we have plotted the "regret" for the waste metric:

$$R_F^A(t) \doteq W_F^A(t) - t\left(b(F) - \sum_j \frac{j}{B} \cdot p_j\right).$$

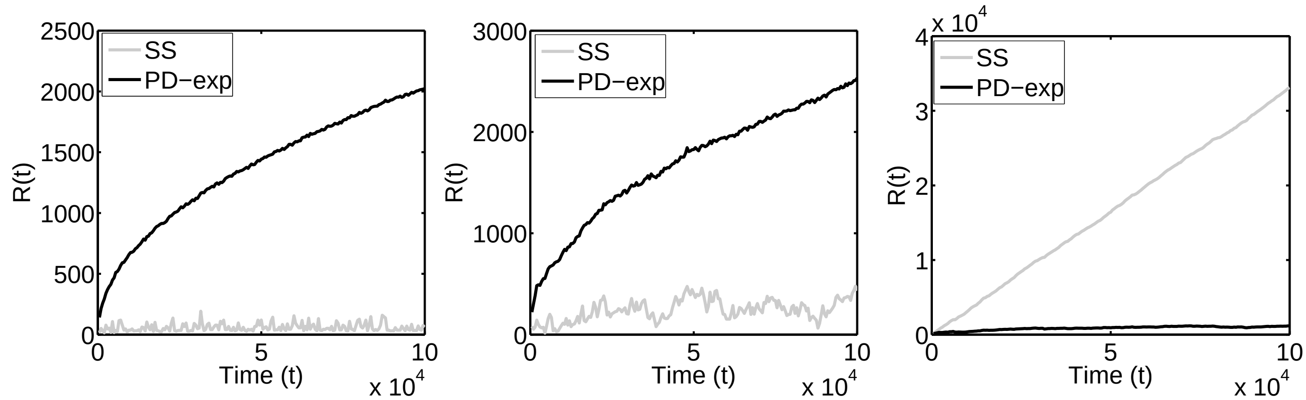
For BW distribution, recall that $R_F^{SS}(t) = W_F^{SS}(t) = \mathcal{O}(\log t)$, whereas our Primal-Dual algorithm only gets a regret $R_F^{PD}(t) = \mathcal{O}(\sqrt{t})$. However, Primal-Dual is still asymptotically optimal for the metric of number of bins used. For PP distribution (middle figure), both SS and Primal-Dual get $\mathcal{O}(\sqrt{t})$ regret, but SS outperforms Primal-Dual. For the simulated LW distribution, the regret of SS is $\Theta(t)$, while it is still $\mathcal{O}(\sqrt{t})$ for Primal-Dual. Therefore, while SS is not asymptotically optimal, Primal-Dual is.

4.2. Performance analysis

We first analyze the case where we know the total number of arrivals T and $\epsilon(t)$ is fixed, and then the more general case when T is not known (*open-ended* bin packing) and $\epsilon(t)$ varies with t. In both cases, the expected number of bins used by the PD-exp algorithm are $\mathcal{O}(\sqrt{T})$ larger than optimal-in-hindsight algorithm.

THEOREM 1. *For i.i.d. item size sequence from distribution F, the PD-exp algorithm with $\epsilon(t) = \sqrt{\frac{B}{T}}$, $\kappa = 1$, and $T > B$, guarantees*

$$\mathbf{E}\left[N_F^{PD}(T)\right] \le T \cdot b(F) + \sqrt{4BT}.$$



(a) BW distribution (b) PP distribution (c) LW distribution

Figure 3 Simulation results comparing performance of SS and PD-exp algorithms for Bounded Waste (BW), Perfectly Packable (PP), and Linear Waste (LW) distributions. The Y-axis shows the difference between algorithms' waste $W_F^A(t)$ and the waste of the optimal LP solution.

THEOREM 2. *For i.i.d. item size sequence from distribution F, the PD-exp algorithm with $\epsilon(t) = \sqrt{\frac{B}{2(B+t)}}$, $\kappa = 1$, guarantees*

$$\mathbf{E}\left[N_F^{PD}(T)\right] \leq T \cdot b(F) + \sqrt{8B(T+B)}.$$

Note that $\mathbf{E}[N_F^{OPT}] \geq T \cdot b(F)$, and therefore the above theorems also imply an expected regret bound compared to optimal offline packing. An identical result also holds for waste $\mathbf{E}[W_F^{PD}(T)]$. The proofs are simple potential function analysis, and appear in Appendix A. A simple application of Azuma-Hoeffding inequality extends our bounds on expected suboptimality to high probability bounds.

THEOREM 3 **(Martingale concentration)**. *For online packing of items from distribution F with horizon T, for all t:*

$$\mathbf{Pr}\left[\sum_h N_h^{PD}(t) \geq t \cdot b(F) + \sqrt{Bt} + \sqrt{BT} + \sqrt{2\lambda\delta t \log t}\right] \leq \frac{1}{t^\lambda}.$$

where $\delta \leq 4 + \frac{\sqrt{B}}{2}$.

4.3. Bounded inventory guarantees

The Primal-Dual algorithm described in Section 4.1 keeps all bins open (potentially $\Theta(T)$ many) for future use, which might be undesirable. We now prove that if instead of $\mathcal{O}(\sqrt{T})$ regret, we only desire $(1+\epsilon)$ competitive ratio, then it is sufficient to keep $\mathcal{O}(1/\epsilon)$ bins open in inventory-at-hand.

DEFINITION 1 (η-BOUNDED INVENTORY ALGORITHMS). Let $\widetilde{N}_h(t)$ denote the number of level h bins open in inventory. An algorithm is an η-bounded inventory algorithm if $\widetilde{N}_h \leq \eta$ for all h. If a level h bin is created at time $t+1$ when $\widetilde{N}_h(t) = \eta$ then the new bin is considered closed. Note that the number of open bins is bounded by ηB at all times.

Algorithm 2 describes the η-bounded inventory algorithm using translated quadratic penalty function.

Algorithm 2 η-bounded **PD-tquad** for one-dimensional level model

---

1: **for** $t = 1, 2, \ldots$ **do**

2: Observe item type Y_t (say, $Y_t = j$)

3: Define level potentials:

$$V_h(t-1) := 1 - \epsilon \left(\eta - \widetilde{N}_h(t-1)\right)^+ \quad (h = 1, \ldots, B-1)$$
$$V_0 := 0 \;; \quad V_B := 1$$

4: Place arriving item in level h^* to create a level $(h^* + j)$ bin, where:

$$h^* = \underset{h:\widetilde{N}_h(t-1)>0}{\arg\min} \; V_{h+j}(t-1) - V_h(t-1)$$

5: If $\widetilde{N}_{h^*+j}(t-1) = \eta$, close one level $(h^* + j)$ bin.

6: **end for**

Theorem 4 **(Bounded inventory)**. *The η-bounded* **PD-tquad** *algorithm $\eta = \frac{1}{\epsilon} + 1$ is asymptotically $(1 + \mathcal{O}(\epsilon))$ competitive.*

The next theorem proves that for a general distribution, any $(1+\epsilon)$ competitive online algorithm must keep $\Omega\left(1/\epsilon\right)$ bins open.

Theorem 5. *For the 1-dimensional bin packing instance with $B = 3$, and items of size 1 and 2 with probability $p_1 = p_2 = \frac{1}{2}$, any online algorithm with a bound of I on the number of open bins must have asymptotic competitive ratio of at least $\left(1 + \frac{1}{9I}\right)$.*

5. Alternate interpretations of Primal-Dual algorithm

In this section we provide some intuitive interpretation of the Primal-Dual algorithms proposed in Section 4.1. This section serves expository purpose; a reader short on time may skip to Section 6 without missing any new result. In Section 5.1 we revisit the Sum-of-squares heuristic and the reason for its failure on Linear Waste distributions. We then show how patching the SS heuristic leads us to the Primal-Dual algorithms with quadratic and translated-quadratic penalty functions. In Section 5.2 we interpret Primal-Dual algorithms as gradient ascent solution of the dual to $\mathbf{P_{1d-level}}$, which is why we name our algorithms Primal-Dual.

5.1. Primal-Dual algorithm as 'patching' the Sum-of-Squares (SS) rule for LW distributions

Recall that SS places arrivals to greedily minimize the penalty function:

$$ss(P) = \sum_{h=1}^{B-1} N_h(P)^2$$

where $N_h(P)$ denotes the number of level h bins in the packing P obtained after packing the arriving item. Therefore, on the arrival of an item of size j at time t, SS sends the item to a bin of level h^* where:

$$h^* = \underset{h:N_h(t-1)>0}{\arg\min} \left[N_{h+j}(t-1) - N_h(t-1)\right]$$

with the convention $N_0 = N_B = 0$. In other words, SS tries to equalize the number of bins at different levels (except level B bins, which are full). This heuristic works for perfectly packable distributions where all items can be packed without any N_h growing large. Thus if any N_h with $h < B$ grows too much, SS stops creating new bins of level h. For example, $B = 5$ and the item size distribution is $p_2 = p_3 = \frac{1}{2}$. In this case, we expect size 2 and size 3 items to be paired together. Now N_4 can not grow a lot because then N_2 must also grow (since SS tries to equalize N_2 and N_4 when packing size 2 items). However, N_2 can not grow too much because size 3 items annihilate them, and no N_h grows as $\Theta(T)$.

However, for Linear Waste distributions some N_h must grow as $\Theta(T)$, and SS heuristic 'pulls along' the number of bins which have room for more items. E.g., for $B = 5$ and only size 2 items, SS creates $\Theta(T)$ bins of level 4 which cause the number of level 2 bins to grow as $\Theta(T)$ as well. Given this intuition, can we modify the SS heuristic to avoid this mode of failure?

One proposal is to stop counting the number of bins once they reach a threshold. In other words, we use the following rule:

$$h^* = \underset{h:N_h(t-1)>0}{\arg\min} \left[(\eta \wedge N_{h+j}) - (\eta \wedge N_h)\right] \tag{4}$$

for a suitable threshold η, and $N_0 = N_B = 0$. It turns out this simple change is not sufficient. Consider the example $B = 8$ with $p_2 = p_5 = \frac{1}{2}$. The optimal packing is to pair size 2 items with a size 5 items. However under the placement rule given above, size 2 items prefer to be packed together to create level 10 bins since level 10 bins are "free." This results in three times as much waste as the offline optimal packing.

Proposal 1: We can fix the above problem by modifying the rule in equation (4) by keeping the convention $N_0 = 0$, but $N_B := \eta$. This makes level B not free. Surprisingly, this heuristic works, and in fact, is identical to the Primal-Dual heuristic with translated quadratic penalty function as we show next.

The penalized-Lagrangian for the translated quadratic penalty function with $\epsilon = \eta$ is:

$$\mathcal{L}_{tquad}(N) = \sum_{h=1}^{B} N_h + \frac{1}{2\eta}\sum_{h=1}^{B-1}\left((\eta - N_h)^+\right)^2.$$

Defining $V_h = \frac{\partial \mathcal{L}}{\partial N_h}$, we get:

$$V_h = \frac{\eta \wedge N_h}{\eta}, \qquad (1 \le h \le B-1);$$
$$V_0 := 0 \ ; \quad V_B := 1.$$

giving the placement rule: $h^* = \arg\min_{h:N_h>0}(V_{h+j} - V_h) = \arg\min_h\left[(\eta \wedge N_{h+j}) - (\eta \wedge N_h)\right]$ with the convention $N_0 = 0, N_B = \eta$. This is precisely Proposal 1.

Proposal 2: Another way to use a sum-of-square potential term while not letting any N_h grow linearly in t is to allow bins to be started at any level h. For example, in the $B = 5$ example with only items of size 2, if we allow bins to start at level 1, then they will reach level 5 and we do not accumulate level 4 bins. This modification by itself is insufficient – nothing prevents us from opening a fresh bin for each arriving item, and packing them at level $B - j$ to create a level B bin (recall there is no N_B^2 term in $ss(P)$). To fix this, we penalize items which cause the level of a bin to reach B. Therefore, on the arrival of tth item of size j, we send it to level h^* where

$$h^* = \arg\min_h\left[C \cdot \mathbf{1}_{h=B-j} + N_{h+j}(t-1) - N_h(t-1)\right] \tag{5}$$

where C is the penalty amount. Note that we do not constrain h^* to the set $\{h : N_h(t-1) > 0\}$. We can equivalently view this heuristic as minimizing the cost function:

$$\widehat{ss}(N) = C \cdot N_B + \frac{1}{2}\sum_{h=1}^{B-1} N_h^2$$

which is a combination of the Best Fit and SS cost functions (but still rather odd due to penalty for 'full bins'). This algorithm "works" and is equivalent to the Primal-Dual heuristic method with quadratic penalty function as we show next, but the output of the Primal-Dual packing must now be viewed upside-down!

For the quadratic penalty function, recall that we must violate all the constraints to get non-zero duals. However, the *(no floating items)* constraints of $\mathbf{P_{1d-level}}$ dictate that for all levels h, there are at least as many items which start at level h than items which end at level h. Violation of these constraints mean that for all h, we must have more items that start at level h than ending at level h. That is, we start packing our bins from top going downwards. Further, since the objective function of $\mathbf{P_{1d-level}}$ is $\sum_j v(j,0)$, we must pay the cost of a bin when an item touches level 0 during this top-down packing.

Denote the violation of level h constraint by Q_h, and let Q_0 denote the number of bins which reach level 0 (we have suppressed the dependent on time t). The penalized-Lagrangian with quadratic penalty function becomes:

$$\mathcal{L}_{quad}(Q) = Q_0 + \frac{\epsilon}{2}\sum_{h=1}^{B-1} Q_h^2$$

and applying the transformation: $N_h = Q_{B-h}$, and $C = \frac{1}{\epsilon}$:

$$= N_B + \frac{1}{2C}\sum_{h=1}^{B-1} N_h^2 = \frac{1}{C}\cdot \widehat{ss}(N).$$

Therefore the actions taken by the Primal-Dual algorithm with quadratic penalty function, and the heuristic in Proposal 2 will be identical.

5.2. Online Mirror Descent view of PD-exp

As the naming of our algorithms suggest, they are in fact Primal-Dual iterations where by the maximization problem that is the dual to $\mathbf{P_{1d-level}}$ is solved using stochastic ascent. We now briefly describe this interpretation of the **PD-exp** algorithm. We start with the LP $\mathbf{P_{1d-level}}$, and perform a change of variables:

$$u(j,h) := \frac{v(j,h)}{p_j}.$$

Note that $u(j) = \{u(j,0), \dots, u(j, B-j)\}$ now represents a probability vector – a randomized strategy for packing item j. Denote the set of feasible action for $u(j)$ by $\mathcal{U}_j = \Delta^{B-j+1}$, the $B-j+1$ dimensional probability simplex. The objective function becomes:

$$\begin{aligned}\sum_j v(j,0) = \sum_{h=1}^{B} n_h &= \sum_{h=1}^{B}\sum_j \left[v(j,h-j) - v(j,h)\right] \\ &= \sum_{h=1}^{B}\sum_j p_j \left[u(j,h-j) - u(j,h)\right]\end{aligned}$$

Introducing duals $\{\alpha_h\}$ $(1 \le h \le B-1)$ for the *(no floating items)* constraints, we write the Lagrangian:

$$\begin{aligned}L(\mathbf{u},\boldsymbol{\alpha}) &= \sum_h \sum_j p_j \left[u(j,h-j) - u(j,h)\right] a - \sum_h \alpha_h \sum_j p_j \left[u(j,h-j) - u(j,h)\right] \\ &= \sum_j p_j \sum_{h=0}^{B-j} u(j,h) \left[(1-\alpha_{h+j}) - (1-\alpha_h))\right]\end{aligned}$$

where we define $\alpha_0 := 1$, $\alpha_B := 0$. This leads to the *dual function*:

$$q(\boldsymbol{\alpha}) := \min_{\{u(j)\in\mathcal{U}_j\}} \sum_j p_j \sum_{h=0}^{B-j} u(j,h)[(1-\alpha_{h+j}) - (1-\alpha_h)] \tag{6}$$

$$= \sum_j p_j \min_{0\le h\le B-j} [(1-\alpha_{h+j}) - (1-\alpha_h)] \tag{7}$$

The dual optimization problem is to solve

$$\boldsymbol{\alpha}^* = \arg\max_{\boldsymbol{\alpha}\ge 0:\alpha_0=1,\alpha_B=0} q(\boldsymbol{\alpha}). \tag{8}$$

The level potentials in Algorithm 1 represent $V_h \equiv 1-\alpha_h$. Let Y_t be the size/type of the tth arrival, and $U_t \in \mathcal{U}_{Y_t}$ denote the action taken by the PD-exp algorithm. Note that

$$N_h(t) = \sum_{s=1}^{t} U_s(h-Y_s) - U_s(h).$$

More importantly, from (6), and noting that in PD-exp items are packed at level h^* that minimizes $(V_{h+j} - V_h)$, we get $\frac{\partial q(\boldsymbol{\alpha})}{\partial \alpha_h} = \mathbf{E}[U_t(h) - U_t(h-Y_t)]$. That, is the primal packing decisions are unbiased stochastic subgradients of the dual function $q(\boldsymbol{\alpha})$.

The Primal-Dual algorithm PD-exp can now be viewed as an Online Mirror Descent (OMD) algorithm (Nemirovski (1979),Beck and Teboulle (2003)) for problem (8) : In a typical application of OMD, a minimization problem $\min_{x\in\Omega} f(x)$ is solved iteratively with access to subgradients $f'(x)$. One first defines a strongly-convex distance function $\omega(x)$ on Ω, and the output $x(t)$ at the tth iteration is set as:

$$x(t) = \arg\min_{x\in\Omega} \left[\epsilon \sum_{s=1}^{t-1} \langle f'(x(s)), x\rangle + \omega(x)\right].$$

The parameter ϵ is called the learning rate.

Applied to our setting, we are solving a maximization problem, and $(U_t(h) - U_t(h-Y_t))$ play the role of $\frac{\partial}{\partial \alpha_h} q(\boldsymbol{\alpha}(t-1))$. By choosing the distance generating function on the space $[0,1]^{B-1}$ as $\omega(\boldsymbol{\alpha}) = \sum_{h=1}^{B-1} \alpha_h (\log \frac{\alpha_h}{\kappa} - 1)$ (strongly convex with respect to ℓ_2 norm), the duals at time t become:

$$\begin{aligned}\alpha_h(t) &= \arg\max_{\alpha\in[0,1]} \left[\epsilon(t) \sum_{s=1}^{t} \langle U_s(h) - U_s(h-Y_s), \alpha\rangle - \alpha\left(\log\frac{\alpha}{\kappa} - 1\right)\right] \\ &= \arg\max_{\alpha\in[0,1]} \left[\epsilon(t)\langle -N_h(t), \alpha\rangle - \alpha\left(\log\frac{\alpha}{\kappa} - 1\right)\right] \\ &= \kappa e^{-\epsilon(t)N_h(t)}.\end{aligned}$$

In other words, the PD-exp algorithm where we greedily minimize the penalized-Lagrangian $\mathcal{L}_{exp}$ is in fact a Primal-Dual iteration folded into a single step.

6. Guarantees against non-$i.i.d.$ input sequences

As we mentioned in the introduction, one of our motivations behind designing distribution-oblivious online packing algorithms is that we hope they will be robust to non-stationary input sequences. For simplicity of presentation, we will focus on the Primal-Dual heuristic with exponential penalty function (Algorithm 1) in this section. How does the PD-exp algorithm perform when the sequence $Y = \{Y_1, \ldots, Y_T\}$ is generated by sampling Y_t from distribution F_t where the sequence $\{F_1, \ldots, F_T\}$ may be arbitrary?

We first formally define the model and notation for non-stationary sequence generation process:
Non-stationary sequence model: We will assume that the item size sequence $Y = \{Y_1, Y_2, \ldots, Y_T\}$ is generated from an arbitrary probability measure μ on $[J]^T$. An equivalent view is the following: At time t, Nature samples an item type Y_t from distribution F_t that is a deterministic function of the history of samples generated: $F_t = f_t(Y_1, \ldots, Y_{t-1})$. Let $\{\mathcal{F}_t\}$ denote the filtration generated by $\{Y_1, \ldots, Y_t\}$ where as is usual $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Therefore, $F_{t|0} \triangleq [F_t|\mathcal{F}_0]$ is a random distribution. Let

$$b_{t|0} = \mathbf{E}\big[b(F_{t|0})|\mathcal{F}_0\big] \ , \qquad w_{t|0} = \mathbf{E}\big[w(F_{t|0})|\mathcal{F}_0\big]$$

denote the expected bins per item and waste per item, respectively, under the random distribution $F_{t|0}$.

Example 1: $Y = \{Y_1,\ Y_2, \ldots, Y_T\}$ is an $i.i.d.$ sequence with distribution F. In this case μ is the product measure $F \times F \cdots \times F$. The random distribution $F_{t|0}$ is degenerate and equals F with probability 1.

Example 2: $Y = \{Y_1,\ Y_2, \ldots, Y_T\}$ is a sequence of independent but not identically distributed item sizes. In this case μ is the product measure $F_1 \times F_2 \cdots \times F_T$. The random distribution $F_{t|0}$ is degenerate and equals F_t with probability 1.

Example 3: Let $Y = \{Y_1, \ldots, Y_T\}$ with $Y_t \in \{2, 3\}$ be generated as follows: define distribution $G = \{p_2 = 1/3, p_3 = 2/3\}$ and distribution $H = \{p_2 = 3/4, p_3 = 1/4\}$. With probability 1/2, nature generates Y from product measure $G \times G \cdots \times G$, and otherwise from product measure $H \times H \cdots \times H$. The random item size distribution $F_{t|0}$ is G with probability 1/2 and H with probability 1/2. Further, $b_{t|0} = \frac{b(G)+b(H)}{2} = \frac{8/27+1/4}{2}$, and $w_{t|0} = \frac{w(G)+w(H)}{2} = 0$.

From the proof of Theorem 1, it is easy to obtain the following corollary:

$$\mathbf{E}\big[N_\mu^{PD}|\mathcal{F}_0\big] \le \sum_{t=1}^{T} b_{t|0} + \mathcal{O}(\sqrt{T}),$$
$$\mathbf{E}\big[W_\mu^{PD}|\mathcal{F}_0\big] \le \sum_{t=1}^{T} w_{t|0} + \mathcal{O}(\sqrt{T}).$$

For example, if Nature generates a non-stationary sequence under the constraint that for all t, the support of F_t is restricted to distributions F with waste $w(F) = 0$, then PD-exp algorithm packs this non-stationary instance with waste $\mathcal{O}(\sqrt{T})$, and hence the regret is $\mathcal{O}(\sqrt{T})$ compared to the optimal offline packing.

However, the above guarantee is rather weak as shown by the following example: Consider the case where $B = 3$, and $Y = \{1, 2, 1, 2, \ldots\}$. In this case $F_{t|0}$ is degenerate distribution $\{p_1 = 1\}$ if t is even, and $\{p_2 = 1\}$ if t is odd. The optimal bin-rate of the former distribution is 1/3 bin/item and for the latter is 1 bin/item, giving an average of 2/3. However, it is conceivable that we should be able to achieve the hindsight optimal bin-rate of 1/2 given the benign nature of input. As we show, we indeed have such a stronger guarantee against a *smoothed* input sequence, albeit at a slightly worse additive suboptimality. In fact, we will prove a family of bounds parametrized by the size of the smoothing window, which hold simultaneously for PD-exp.

To present our results on guarantees against smoothed instances, we need to define what we call *L-bounded fractional partition* of the time horizon $[T]$.

DEFINITION 2 (L-BOUNDED FRACTIONAL PARTITION). A collection $\pi = \{q_1, q_2, \ldots, q_T\}$ of T probability measures, each with support on $[T]$ is called an L-bounded fractional partition of the time horizon $\{1, 2, \ldots, T\}$ if it satisfies:

1. **Bounded L-Support:** For all $k \in [T]$, $\max\{t : q_k(t) > 0\} - \min\{t : q_k(t) > 0\} \le L - 1$
2. **Unit marginals:** For all $t \in [T]$: $\sum_k q_k(t) = 1$

Let $\ell_k = \min\{t : q_k(t) > 0\}$ denote the earliest time t in the support of kth window, and without loss of generality assume the windows are indexed so that ℓ_k is non-decreasing in k. Let $\Pi(L, T)$ denote the set of all L-bounded fractional partitions of $[T]$.

See Figure 4 for an illustration. In words, window k is defined by a probability measure q_k on $\{1, \ldots, T\}$ which gives the mixing weights, and these mixing weights give a smoothed item size distribution. The L-boundedness condition says that no smoothing window can have support on time instants more than L apart, and the distribution for each time instant t is completely partitioned across the T windows in π. We will assume that the partition is deterministic, that is independent of $\{Y_1, \ldots, Y_T\}$. Next, we make the notion of smoothed distribution formal.

DEFINITION 3. Given an L-bounded fractional partition $\pi = \{q_k\}_{k \in [T]}$, and the adversarial model of item type generation described above, the kth *smoothed item size distribution* conditioned on $\mathcal{F}_t$ is defined to be:

$$\hat{F}_{\pi,k|t} \triangleq \mathbf{E}\left[\sum_{u=1}^{T} q_k(u) F_u \,\middle|\, \mathcal{F}_t \right] \tag{9}$$

Denote the optimal bin rate of $\mathbf{P_{1d-level}}$ with distribution $\hat{F}_{\pi,k|t}$ by $\hat{b}_{\pi,k|t}$. For succinctness, $\hat{b}_{\pi,k} \triangleq \hat{b}_{\pi,k|0}$. Note that $\hat{F}_{\pi,k|t}$ and $\hat{b}_{\pi,k|t}$ are $\mathcal{F}_t$-adapted random elements.

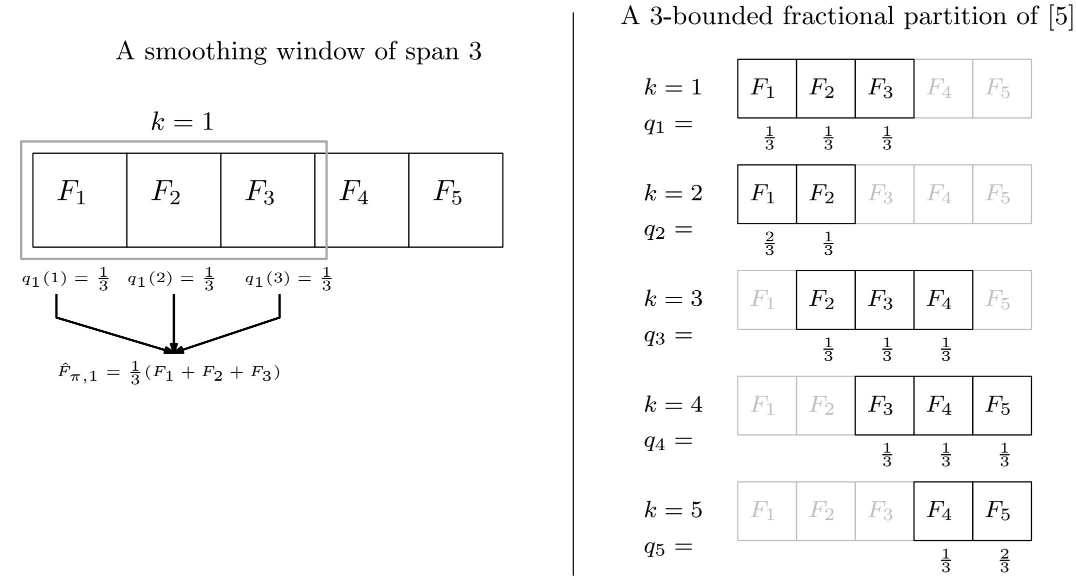


Figure 4 An illustration of L-bounded fractional partitions of $[T] = \{1, 2, \ldots, T\}$.

We are now ready to state our main result of this section.

THEOREM 6. *Consider the non-stationary arrival model described above with a packing horizon of T items, and an arbitrary conditioning time t. Let ϵ and κ denote the parameters of the penalized-Lagrangian function in Algorithm 1.*

If L, ϵ, κ, b satisfy: $\epsilon L < 1$, and $L \leq \frac{1}{\epsilon} \log \frac{\kappa B}{B-1}$, then for all $\pi \in \Pi(L, T)$:

$$\mathbf{E}\left[N_\mu^{PD}(T) | \mathcal{F}_t\right] \leq \sum_{k=1}^{T} \hat{b}_{\pi,k|t} + T\kappa\epsilon \left(\frac{2L+3}{4}\right) + \frac{B\kappa}{\epsilon}.$$

The proof of Theorem 6 appears in Appendix A. The main idea is to calculate the change in the Lagrangian over a random time distributed according to the measure q_k. We can view this as calculating the change in Lagrangian in a single time step with item size distribution being the smoothed distribution $\hat{F}_{\pi,k|t}$ but where the packing seen by the arriving job is random, and potentially correlated with the type of the arriving item. Since the support of the random time is bounded by L, the support of the random packing is "narrow". This allows us a graceful increase in suboptimality.

Note that Theorem 6 gives a family of bounds: For any setting of ϵ, as the smoothing window bound L is increased from $\mathcal{O}(1)$ to $o(1/\epsilon)$ the first term $\sum_k \hat{b}_{\pi,k|t}$ is non-decreasing since the objective function $b(F)$ in $\mathbf{P}_{\mathbf{1d-level}}$ is a convex function of F. However, the additive regret term becomes larger. Since the theorem is too general to be directly applicable, we give some concrete corollaries next.

Corollary 1. *If $\epsilon = \Theta(T^{-1/2})$ is fixed, then for $L = O(T^c)$ with any $c < \frac{1}{2}$:*

$$\mathbf{E}\big[N_\mu^{PD}(T)\big] \leq \min_{\pi \in \Pi(L,T)} \left[\sum_{k=1}^{T} \hat{b}_{\pi,k}\right] + O(T^{\frac{1}{2}+c}).$$

Corollary 2. *If the size of the desired smoothing window is given as $L = \Theta(T^c)$ $(0 \leq c < 1)$, then choosing $\epsilon = \Theta(T^{-\frac{1+c}{2}})$:*

$$\mathbf{E}\big[N_\mu^{PD}(T)\big] \leq \min_{\pi \in \Pi(L,T)} \left[\sum_{k=1}^{T} \hat{b}_{\pi,k}\right] + O(T^{\frac{1+c}{2}}).$$

Corollary 1 again gives a family of bounds parametrized by L for a fixed ϵ, while Corollary 2 dictates an optimal choice of ϵ given L to minimize the regret term. corollary 1 implies that for the input sequence $Y = \{1,2,1,2,\cdots\}$, the PD-exp algorithm as-is gives an $\mathcal{O}(\sqrt{T})$ regret compared to offline optimal packing.

The following is a further application of the above results to *almost-stationary sequences.*

Corollary 3 **(Local adversarially perturbed input)**. *Let $Y = \{Y_1, \ldots, Y_T\}$ be an i.i.d. sequence of item types generated from F. Consider an adversary who applies a permutation σ (possibly a function of Y) to Y before presenting $\tilde{Y} = \{Y_{\sigma(1)}, \ldots, Y_{\sigma(T)}\}$ to the online packing algorithm. We assume the adversary is constrained in his choice of σ as follows: $\forall k,\ k - \frac{W}{2} \leq \sigma(k) \leq k + \frac{W}{2}$.*

1. *If $W = o(\sqrt{T})$ is unknown, then setting $\epsilon = \Theta(T^{-1/2})$ (that is, $L = T^{1/4}W^{1/2}$)*

$$\mathbf{E}\big[N^{PD}(T)\big] \leq b(F) \cdot T + \mathcal{O}(T^{3/4}W^{1/2}).$$

2. *If $W = o(T)$ is known, then setting $\epsilon = \Theta(T^{-2/3}W^{-1/3})$ (that is, $L = T^{1/3}W^{2/3}$):*

$$\mathbf{E}\big[N^{PD}(T)\big] \leq b(F) \cdot T + \mathcal{O}(T^{2/3}W^{1/3}).$$

Proof: For the first part of the corollary, we apply Corollary 1 to a simple L-bounded fractional partition π^* where $q_{iL+1} = q_{iL+2} = \cdots = q_{(i+1)L} = \text{Unif}[\{iL+1, \ldots, (i+1)L\}]$ for $i = 1, \ldots, \lfloor \frac{T}{L} \rfloor$. That is, we partition the horizon into disjoint contiguous windows of length L, and our main goal is to optimize L for use in the analysis (note the algorithm only depends on ϵ which we have fixed to $\Theta(T^{-1/2})$). Now for any k, the total variation norm of the smoothed distribution $\hat{F}_{\pi^*,k}$ and the true distribution F is $\mathcal{O}(W/L)$ and therefore $\mathbf{E}\left[\hat{b}_{\pi^*,k}\right] = b(F) + \mathcal{O}(W/L)$. Applying Corollary 1, we get

$$\mathbf{E}\big[N^{PD}(T)\big] \leq b(F) \cdot T + \mathcal{O}\left(\frac{WT}{L} + \frac{TL}{\epsilon} + \frac{1}{\epsilon}\right)$$

This expression is minimized for $L = \Theta(T^{1/4}W^{1/2})$ giving the regret bound in the corollary statement.

Proof of the second case is similar, except when we optimize for L, we set $\epsilon = \Theta\left(\sqrt{\frac{1}{TL}}\right)$ as per Corollary 2. □

Remark 1. Our analysis in Corollary 3 is likely not tight. For $W = 1$, the adversary is restricted to swapping some subset of adjacent pairs of the original sequence (after having seen the sequence). It would be surprising if this small perturbation causes the regret to jump from $\mathcal{O}(\sqrt{T})$ to $\mathcal{O}(T^{3/4})$. A likely line of attack might be to revert to the Online Mirror Descent (OMD) view of PD-exp (Section 5.2), and then develop tighter regret analysis of OMD under a general correlation structure on gradients. While some analysis exists under mixing assumptions on the input sequence (Agarwal and Duchi (2013)), adversarial perturbations of *i.i.d.* sequences do not belong to this class of inputs.

In Section 8 we conjecture the existence of a complementary non-stationary guarantee to the one proved in this Section: If each F_t lies in a certain set (Corner-Point distributions with a common unique optimal dual solution), then setting $\epsilon = \Theta(\log^{1+\beta} T)$ for any $\beta > 0$ will result in expected regret $\mathcal{O}(\log^{1+\beta} T)$ against the optimal offline packing.

7. Configuration Model for Bin Packing

In this section we present generalization of the one-dimensional model described in Section 2 to configuration model. The flexibility of configuration model allows modeling multi-dimensional bin packing and vector packing problems, as well as minimum cost packing. We briefly define the configuration model in Section 7.1 and the extension of our algorithm in this setting in Section 7.2.

7.1. Model definitions

As before, item types Y_t are *i.i.d.* samples from $[J] = \{1, \ldots, J\}$, with p_j denoting the probability of type j. We abbreviate $F = \{p_1, \ldots, p_J\}$. Unlike one-dimensional level model where packing constraints were implicitly dictated by bin capacity and item sizes, here the set of feasible configurations is explicitly defined and denoted by $\mathcal{C}$. Each $c \in \mathcal{C}$ is represented by $x_c = \{x_{c1}, \ldots, x_{cJ}\} \in \mathbb{Z}^J$. The quantity x_{cj} represents the number of type j items in configuration c. We assume that the set $\mathcal{C}$ is downward closed: if $c \in \mathcal{C}$, then all c' with $x_{c'j} \leq x_{cj}$ $\forall j$ are in $\mathcal{C}$ as well. (The model can be extended to allow heterogeneous bins; we omit it for brevity). We denote the empty configuration by $\emptyset$, the configuration with a single type j item by e_j, and $c + e_j$ and $c - e_j$ denote configurations with one more and one less item j, respectively (assuming they are in $\mathcal{C}$). We denote the number of configuration c bins in packing P by $N_c(P)$, and with abuse of notation $N_c^A(t)$ will denote the number of bins in configuration c after tth item has been packed by algorithm A.

The performance metric that we want to minimize is the number of bins opened by A :

$$N(P) \triangleq \sum_{c \in \mathcal{C} \setminus \emptyset} N_c(P).$$

For the minimum cost packing problem, we associate a weight w_c to configuration c bins and define the weight of a packing P by:

$$W(P) \triangleq \sum_{c \in \mathcal{C}} w_c N_c(P).$$

As earlier, $P_F^A(T)$ denotes the packing after applying algorithm A to T items generated from F, and the expected costs $\mathbf{E}[W_F^A(T)]$ and $\mathbf{E}[N_F^A(T)]$ are defined analogously to (1)-(2). Note that the minimization problems for $N(P)$ and $W(P)$ are not equivalent unless w_c satisfy: $c = qc_1 + (1-q)c_2$ implies $w_c = qw_{c_1} + (1-q)w_{c_2}$. To keep exposition simple, we will focus on the metric of minimizing total number of bins in this section.

7.2. Bin packing Linear Program

Analogous to $\mathbf{P_{1d-level}}$, we begin with a Linear Program that given the item type distribution F finds the optimal number of bins per item for F:

$$\begin{aligned} b(F) = &\min_{\{n_c\}_\mathcal{C}} \sum_{c \in \mathcal{C}} n_c && (\mathbf{P_{conf}}) \\ \text{subject to} \quad &\forall c \in \mathcal{C} \;:\; n_c \geq 0 \\ &\forall j \in [J] \;:\; \sum_c n_c x_{cj} = p_j \end{aligned}$$

The variables n_c denote the expected number of configuration c bins opened per item from F, and the non-negativity constraints ($n_c \geq 0$) are analogous to *(no floating items)* constraints in $\mathbf{P_{1d-level}}$. The second set of constraints are the mass balance constraints.

7.3. Primal-Dual algorithm

Starting from $\mathbf{P_{conf}}$ the penalized Lagrangian where we dualize the $n_c > 0$ constraints using exponential penalty function is given by:

$$\mathcal{L}_{exp}(n) = \sum_c n_c + \kappa\epsilon \sum_c e^{-n_c/\epsilon}.$$

As in Section 4.1, $n_c \sim \frac{N_c(t)}{t}$, giving:

$$\mathcal{L}_{exp}(N(t)) = \sum_c N_c(t) + \frac{\kappa}{\epsilon(t)} \sum_c e^{-\epsilon(t) N_c(t)}.$$

Finally, we obtain Algorithm 3 with the following performance guarantee. We omit the proof since it is very similar to the proof of Theorem 2.

Theorem 7. *For i.i.d. item type sequences from a distribution F, the PD-exp algorithm with $\epsilon(t) = \sqrt{\frac{|\mathcal{C}|}{2(|\mathcal{C}|+t)}}$, $\kappa = 1$, guarantees:*

$$\mathbf{E}\left[N_F^{PD}(T)\right] \leq \mathbf{E}\left[N_F^{OPT}(T)\right] + \sqrt{8|\mathcal{C}|(T+|\mathcal{C}|)}.$$

Algorithm 3 PD-exp for configuration model

1: **for** $t = 1, 2, \ldots$ **do**

2: Observe item type Y_t (say, $Y_t = j$)

3: Define configuration potentials:

$$V_c(t-1) := 1 - \kappa e^{-\epsilon(t) N_c(t-1)}$$
$$V_\emptyset := 0$$

4: Place arriving item in configuration c^* to create $c^* + e_j$, where:

$$c^* = \underset{c: N_c(t-1) > 0}{\arg\min}\; V_{c+e_j}(t-1) - V_c(t-1)$$

5: **end for**

8. Beyond Bounded Waste and Linear Waste: Corner Point distributions

The Courcoubetis and Weber (1986) characterization of distributions as Linear Waste, Perfectly Packable, and Bounded Waste seems discouraging: most distributions have Linear Waste which are bad instances for online packing. In this section we conjecture that from an algorithm design viewpoint, the opposite situation holds. We provide evidence that most distributions are in fact easy to pack; the Primal-Dual Algorithm 3 can pack them with $\mathcal{O}(\log^{1+\beta} T)$ regret (for any $\beta > 0$) compared to the optimal in hindsight. Further, this $\mathcal{O}(\log^{1+\beta} T)$ regret result is robust to non-stationary input sequences under the constraint that each distribution in the sequence is at most a small perturbation from a base distribution.

To make a more formal conjecture, it is more convenient to use the configuration model from Section 7.2. We begin with the dual for problem $\mathbf{P_{conf}}$ by introducing dual variables α_c for the first set of non-negativity constraints, and z_j for the second set of mass conservation constraints:

$$\begin{aligned} &\max_{\{z_j\}_{[J]}, \{\alpha_c\}_{\mathcal{C}}} \sum_{j \in [J]} p_j z_j && (\mathbf{D_{conf}}) \\ \text{subject to} \quad & \forall c \in \mathcal{C} \;:\; \alpha_c \geq 0 \\ & \forall c \in \mathcal{C} \;:\; \sum_j z_j x_{cj} = 1 - \alpha_c \end{aligned}$$

which we further simplify to:

$$\begin{aligned} &\max \sum_{j \in [J]} p_j z_j && (\mathbf{D'_{conf}}) \\ \text{subject to} \quad & \forall c \in \mathcal{C} \;:\; \sum_j z_j x_{cj} \leq 1 \end{aligned}$$

The $|\mathcal{C}|$ constraints define a polytope of feasibility for $\{z_j\}$, call it $\mathcal{Z}$. Note that $\mathcal{Z}$ is a function of the support of the distribution U_F and the set of feasible configurations $\mathcal{C}$. Let $z^*(F) \in \mathcal{Z}$ be a maximizer of $\mathbf{D'_{conf}}$ for distribution F.

Definition 4 (Corner-Point distribution). A distribution F is called a corner point (CP) distribution if $z^*(F)$ is unique. Note that $z^*(F)$ must then necessarily be an extreme point of the polytope $\mathcal{Z}$.

CP distributions are universal in the sense that for a given support U of item types, any perturbation of a distribution (with no atoms in the perturbing noise) gives a CP distribution with probability 1. Further, they are robust in the sense that if we start with a CP distribution, then there is a small enough perturbation keeps the distribution CP with the same corner z^*.

For an extreme point $z^\dagger$ of $\mathcal{Z}$, let $\mathcal{P}(z^\dagger)$ denote the set of all CP distributions corresponding to $z^\dagger$:

$$\mathcal{P}(z^\dagger) = \{F | z^*(F) = z^\dagger \text{ is unique}\}.$$

Let $\mathcal{P}(z^\dagger, \delta)$ denote the set of all CP distributions whose δ ball is contained in $\mathcal{P}(z^\dagger)$:

$$\mathcal{P}(z^\dagger, \delta) = \{F | F' \in \mathcal{P}(z^\dagger) \ \forall \ ||F' - F||_\infty \leq \delta\}.$$

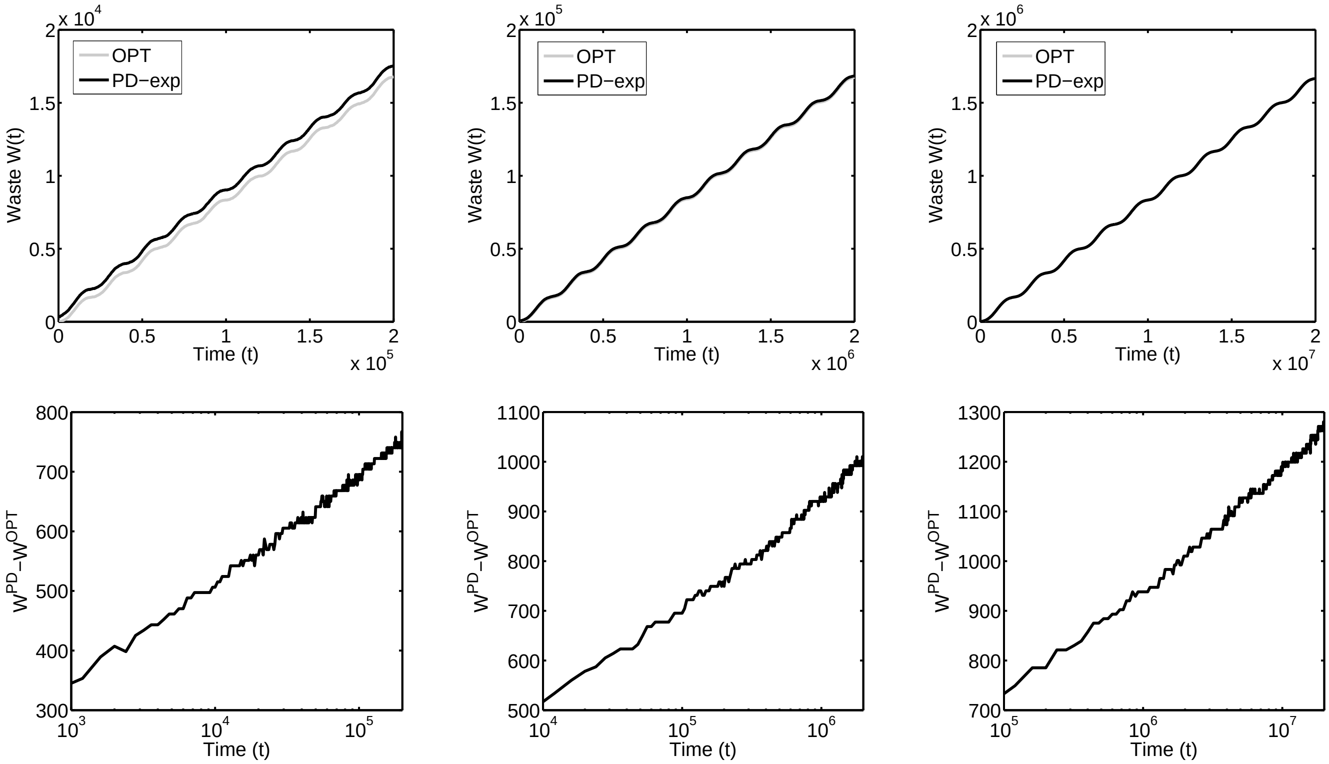
We next present our conjecture which states that CP distributions are in fact easy to pack online.

Conjecture 1 (**$\mathcal{O}(\log^{1+\beta} T)$ regret for CP distributions**). *Let the packing instance $Y = \{Y_1, \ldots, Y_T\}$ be generated by sampling Y_t from F_t independently, and where each $F_t \in \mathcal{P}(z^\dagger, \delta)$ for some $\delta > 0$, and some extreme point $z^\dagger$ of $\mathcal{Z}$. Then, applying the PD-exp Algorithm 3 with $\epsilon(t) = \Theta\left(\frac{1}{\log^{1+\beta} t}\right)$ for any $\beta > 0$,*

$$\mathbf{E}\big[N^{PD}(T) - N^{OPT}(T)\big] \leq \mathcal{O}(\log^{1+\beta} T).$$

Conjecture 1 is complementary to Theorem 6 : If for an L-bounded fractional partition π, the smoothed distributions $\{\hat{F}_{\pi,1|0}, \hat{F}_{\pi,2|0}, \ldots\}$ satisfy the conditions given above, then we conjecture that an appropriate choice of $\epsilon(t)$ in PD-exp algorithm should achieve a regret that grows polylogarithmically in T, and at most polynomially in L.

Before we provide intuition for why we expect Conjecture 1 to be true, we provide some experimental evidence. In Figure 5, we show three single simulation run of a non-stationary online packing instance with $T = 2 \times 10^5$, $T = 2 \times 10^6$, and 2×10^7. The bin size is $B = 9$, and the support of the distributions is $U = \{2, 3\}$ The distribution at time t is given by a convex combination of $F_L = \{p_2 = \frac{43}{48}, p_3 = \frac{5}{48}\}$ and $F_U = \{p_2 = \frac{37}{48}, p_3 = \frac{11}{48}\}$. The weights are given by a sinusoidal function whose period was chosen to accommodate 10 full cycles of the sin function. All the distributions thus used are CP distributions associated with the unique dual $z = \{z_2 = \frac{1}{4}, z_3 = \frac{1}{4}\}$. We set $\epsilon(t) = \frac{1}{\log^{1.5}(t+10)}$ in Algorithm 3. The waste of PD-exp is remarkably close to the offline optimal and, as shown in the second plot, the regret appears to grow polynomially in $\log t$.



(a) $T = 2 \times 10^5$ (b) $T = 2 \times 10^6$ (c) $T = 2 \times 10^7$

Figure 5 Simulation results showing the performance of PD-exp on a non-stationary arrival sequence where the distributions F_t are all CP distribution with a common unique dual. The arrival distributions were generated by taking convex combination of two distributions, with weights given by a sinusoidal function. The period of the sin function was chosen to accommodate 10 cycles in the total number of arrivals simulated. The top subplots show waste of the PD-exp algorithm and the waste of the optimal offline (which can repack items) vs. number of arrivals. The bottom subplot shows the regret vs. time (log-scale). The growth rate of regret in each subplot, as well the final values across the three subplots suggest that the regret is growing as $\mathcal{O}(\text{polylog}\, t)$.

Intuition for Conjecture 1 : First observe that for a CP distribution, given the unique optimal dual solution z^*, one optimal solution for dual variables for the non-negativity constraints $n_c \geq 0$ is given by: $\alpha^*_c = 1 - \sum_j z^*_j x_{cj}$. By complementary slackness, $\alpha^*_c > 0$ implies $n_c = 0$ in the optimal solution. Call the configurations c with $\alpha^*_c > 0$ non-basic configurations, and the rest basic configurations. To prove the conjecture, it would suffice if we prove that $N_c(T) = \mathcal{O}(\log^{1+\beta} T)$ for all non-basic configurations. We now use the observation that the Primal-Dual algorithm is essentially a stochastic gradient ascent algorithm for maximizing the dual function $q(\alpha)$ (analogous to (6)), of which α^* is a maximizer. Further, for all distributions within $\mathcal{P}(z,\delta)$, the optimal solution is the same. Now we would be done if the following two conditions were met: (i) α^* is unique, and (ii) the subgradient of function $q(\alpha)$ is bounded away from 0 everywhere. This would imply that $\mathbf{E}\left[e^{\gamma||N(t)-N^*(t)||}\right] = \mathcal{O}(1)$ (for some constant $\gamma > 0$) where $N^*(t)$ is the state corresponding to optimal duals α^*, and therefore, $\mathbf{Pr}[||N(t) - N^*(t)|| \geq K \cdot \log T] = o(1/T)$ (for some constant $K > 0$).

However, in our case α^* is not unique, and therefore $q(\alpha)$ does not possess subgradients bounded away from 0 everywhere. This is the challenge in proving the Conjecture. Fortunately, an analysis along similar lines as that of BW distributions in Csirik et al. (2006) seems possible – we further classify the non-basic configurations into *live configurations* (those to which items can be added to reach a basic configuration), and *dead-end configurations* (all others). We now observe that α_c^* is unique for all live configurations c, and α_c^* for dead-end configurations lie in a low dimensional subspace. The subgradient of the dual function $q(\alpha)$ are bounded away from 0 when restricted to deviations from these subspaces.

Conjecture 1 is reminiscent of recent results on network revenue management where uniqueness of the dual solution to the offline LP is used to prove that resolving heuristics with probabilistic acceptance rules give $\mathcal{O}(1)$ regret in the case of known demand parameters (Jasin and Kumar (2012)), and $\mathcal{O}(\log^2 T)$ regret in the case of unknown demand parameters which are learned online (Jasin (2015)). Our conjecture suggests that in the case of unknown demand parameters, it may be possible to obtain $\mathcal{O}(\log^{1+\beta} T)$ regret without any explicit learning, even when demand is somewhat non-stationary.

9. Summary and Open Questions

Our driving goal in this paper was to design simple, distribution-oblivious and asymptotically optimal algorithms for online stochastic packing. We proved that the bin packing Linear Program can be turned into online packing algorithms using penalized Lagrangian similar to interior-point methods which achieve $\mathcal{O}(\sqrt{T})$ regret for all distributions. Our algorithms can also be viewed as Lyapunov function based control algorithms, and indeed we leverage this interpretation to prove guarantees for non-stationary inputs, which appear to be novel to the best of our knowledge. At the same time, our guarantees are quite naive, and tightening the upper bounds is an interesting line of future work.

We also introduced the notion of Corner Point (CP) distributions – a class to which almost all distributions belong. One question that we leave open is proving Conjecture 1, that Corner Point distributions can be packed with $\mathcal{O}(\log^{1+\beta} T)$ regret by our Primal-Dual heuristic.

Open Questions:

Assuming Conjecture 1, our distribution oblivious algorithms still require the knowledge of whether the input distribution has a unique dual or not to get the improved $\mathcal{O}(\text{polylog}\, T)$ regret for corner point distributions. Whether there are algorithms which can without any tunable parameters simultaneously achieve $\mathcal{O}(\text{polylog}\, T)$ regret for CP distributions, and $\mathcal{O}(\sqrt{T})$ regret for non-CP distributions is an open question.

Given the promising performance of our distribution-oblivious heuristics, one natural question is *what is the cost of obliviousness, or the value of learning*? We adopted a philosophical stance of distribution-obliviousness because the resulting algorithms are simple, and turn out to be robust – more so with the notion of CP distributions in hand. Exploring the best marriage of learning algorithms with Lyapunov function based control is an interesting avenue for future research.

A final open question is exploring alternate notions of non-stationary instances: There is substantial literature on bin packing under adversarial sequences, as well as for *i.i.d.* input sequences. In literature on control of queueing systems and in machine learning, researchers have started exploring input sequences which come from a Markov chain, or which exhibit mixing, so the auto-correlation function decays exponentially fast. One flavor of result we have presented in this paper is for locally perturbed *i.i.d.* sequences. Two questions to explore here are: *(a)* tightening our regret analysis, and *(b)* whether locally perturbed *i.i.d.* sequences are strictly harder than Markov/mixing input sequences (in terms of regret). However our results also suggest another alternate to analyzing non-stationary instances – rather than varying the strength of the instance generating process, we vary the strength of the benchmark we compare the performance against. In this vain we presented a regret result with respect to sum of LP solutions of local smoothing of the input sequence. A challenging open question to explore here is whether similar regret results can be obtained against online packing oracles which can look ahead some finite time steps into the future.

References

Adelman, Daniel, George L. Nemhauser. 1999. Price-directed control of remnant inventory systems. *Operations Research* **47**(6) 889–898.

Agarwal, Alekh, John C Duchi. 2013. The generalization ability of online algorithms for dependent data. *Information Theory, IEEE Transactions on* **59**(1) 573–587.

Agrawal, Shipra, Nikhil R Devanur. 2015. Fast algorithms for online stochastic convex programming. *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 1405–1424.

Albers, Susanne, Michael Mitzenmacher. 1998. Average-case analyses of first fit and random fit bin packing. *SODA '98: Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 290–299.

Awerbuch, Baruch, Rohit Khandekar. 2008. Stateless distributed gradient descent for positive linear programs. *Proceedings of the 40th annual ACM symposium on Theory of computing*. STOC '08, ACM, New York, NY, USA, 691–700. doi:10.1145/1374376.1374476. URL http://doi.acm.org/10.1145/1374376.1374476.

Beck, Amir, Marc Teboulle. 2003. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters* **31**(3) 167–175.

Boyd, Stephen, Lieven Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, New York, NY, USA.

Buchbinder, Niv, Kamal Jain, Joseph Seffi Naor. 2007. Online primal-dual algorithms for maximizing ad-auctions revenue. *Proceedings of the 15th annual European conference on Algorithms*. ESA'07, Springer-Verlag, Berlin, Heidelberg, 253–264. URL http://dl.acm.org/citation.cfm?id=1778580.1778606.

Coffman, E. G., Jr., C. Courcoubetis, M. R. Garey, D. S. Johnson, L. A. McGeoch, P. W. Shor, R. R. Weber, M. Yannakakis. 1991. Fundamental discrepancies between average-case analyses under discrete and continuous distributions: a bin packing case study. *STOC '91*. ACM, New York, NY, USA, 230–240. doi:http://doi.acm.org/10.1145/103418.103446.

Coffman, E.G., A.G. Stolyar. 2001. Bandwidth packing. *Algorithmica* **29**(1-2) 70–88.

Courcoubetis, C, RR Weber. 1986. Necessary and sufficient conditions for stability of a bin-packing system. *Journal of applied probability* 989–999.

Csirik, János, David S. Johnson, Claire Kenyon, James B. Orlin, Peter W. Shor, Richard R. Weber. 2006. On the sum-of-squares algorithm for bin packing. *J. ACM* **53**(1) 1–65.

Csirik, János, David S. Johnson, Claire Kenyon, Peter W. Shor, Richard R. Weber. 1999. A self organizing bin packing heuristic. *Selected papers from the International Workshop on Algorithm Engineering and Experimentation*. ALENEX '99, Springer-Verlag, London, UK, 246–265. URL http://portal.acm.org/citation.cfm?id=646678.702165.

Gamarnik, David. 2004. Stochastic bandwidth packing process: Stability conditions via lyapunov function technique. *Queueing Syst. Theory Appl.* **48**(3-4) 339–363. doi:http://dx.doi.org/10.1023/B:QUES.0000046581.34849.cf.

Gamarnik, David, Mark S. Squillante. 2005. Analysis of stochastic online bin packing processes. *Stochastic Models* **21** 401 – 425.

Gupta, Anupam, Marco Molinaro. 2014. How experts can solve lps online. *Algorithms-ESA 2014*. Springer, 517–529.

Iyengar, Garud, Karl Sigman. 2004. Exponential penalty function control of loss networks. *Annals of Applied Probability* 1698–1740.

Jasin, Stefanus. 2015. Performance of an lp-based control for revenue management with unknown demand parameters. *Operations Research* **63**(4) 909–915.

Jasin, Stefanus, Sunil Kumar. 2012. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research* **37**(2) 313–345.

Johnson, David S., Alan Demers, Jeffrey D. Ullman, Michael R Garey, Ronald L. Graham. 1974. Worst-case performance bounds for simple one-dimensional packing algorithms. *SIAM Journal on Computing* **3**(4) 299–325.

Karmarkar, Narendra, Richard M. Karp. 1982. An efficient approximation scheme for the one-dimensional bin-packing problem. *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*. SFCS '82, IEEE Computer Society, Washington, DC, USA, 312–320. doi:http://dx.doi.org/10.1109/SFCS.1982.61. URL http://dx.doi.org/10.1109/SFCS.1982.61.

Kenyon, C., M. Mitzenmacher. 2000. Linear waste of best fit bin packing on skewed distributions. *Proceedings. 41st Annual Symposium on Foundations of Computer Science*. 582 –589. doi:10.1109/SFCS.2000.892326.

Kenyon, Claire, Yuval Rabani, Alistair Sinclair. 1996. Biased random walks, lyapunov functions, and stochastic analysis of best fit bin packing. *SODA '96: Proceedings of the seventh annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 351–358.

Leighton, F. T., P. Shor. 1986. Tight bounds for minimax grid matching, with applications to the average case analysis of algorithms. *STOC '86: Proceedings of the eighteenth annual ACM symposium on Theory of computing*. ACM, New York, NY, USA, 91–103. doi:http://doi.acm.org/10.1145/12130.12140.

Lelarge, Marc. 2007. Online bandwidth packing with symmetric distributions. *AofA'07*. 471–482.

Nemirovski, A. 1979. Efficient methods for large-scale convex optimization problems. *Ekonomika i Matematicheskie Metody* **15**.

Plotkin, Serge A., David B. Shmoys, Éva Tardos. 1995. Fast approximation algorithms for fractional packing and covering problems. *Mathematics of Operations Research* **20**(2) pp. 257–301. URL http://www.jstor.org/stable/3690406.

Rhee, WanSoo T., Michel Talagrand. 1993. On-line bin packing of items of random sizes, II. *SIAM J. Comput.* **22** 1251–1256.

Rothvoß, Thomas. 2013. Approximating bin packing within o (log opt* log log opt) bins. *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 20–29.

Shah, Devavrat, John N. Tsitsiklis. 2008. Bin packing with queues. *J. Appl. Prob.* **45**(4) 922–939.

Shor, P. W. 1986. The average-case analysis of some on-line algorithms for bin packing. *Combinatorica* **6**(2) 179–200. doi:http://dx.doi.org/10.1007/BF02579171.

Shor, P.W. 1991. How to pack better than best fit: tight bounds for average-case online bin packing. *Foundations of Computer Science, 1991. Proceedings., 32nd Annual Symposium on*. 752 –759. doi:10.1109/SFCS.1991.185444.

Stolyar, Alexander L. 2005. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Syst. Theory Appl.* **50**(4) 401–457. doi:10.1007/s11134-005-1450-0. URL http://dx.doi.org/10.1007/s11134-005-1450-0.

Tassiulas, L., A. Ephremides. 1992. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *Automatic Control, IEEE Transactions on* **37**(12) 1936–1948.

Endnotes

1. The parameters of the three distributions are as follows:

BW distribution: $B = 9,\ F = \{p_2 = \frac{35}{48}, p_3 = \frac{13}{48}\}$

PP distribution: $B = 10,\ F = \{p_1 = \frac{1}{4}, p_3 = \frac{1}{4}, p_4 = \frac{1}{8}, p_5 = \frac{1}{4}, p_8 = \frac{1}{8}\}$

LW distribution: $B = 10,\ F = \{p_3 = \frac{1}{4}, p_4 = \frac{1}{4}, p_5 = \frac{1}{4}, p_8 = \frac{1}{4}\}$

Appendix A: Proofs

A.1. Proof of Theorem 1

The penalized-Lagrangian is:

$$\mathcal{L}(\mathbf{N}(t)) = \underbrace{\sum_{h=1}^{B} N_h(t)}_{F_A(t)} + \underbrace{\frac{\kappa}{\epsilon}\sum_{h=1}^{B-1} e^{-\epsilon N_h(t)}}_{V_A(t)} \tag{10}$$

The change in Lagrangian on adding an item j to level h_1 to create $h_2 = h_1 + j$ is:

$$\Delta\mathcal{L}(\mathbf{N}) \leq \mathbf{1}_{\{h_1=0\}} + \frac{\kappa}{\epsilon}\sum_{h} e^{-\epsilon N_h}\left(e^{-\epsilon\cdot\mathrm{sgn}(\Delta N_h)} - 1\right)$$

where N_h denote the number of bins of respective levels before the move.

We now devise a distribution aware probabilistic packing policy A_F and then upper bound the increase in the Lagrangian under PD heuristic (for a given state during the PD packing) by that under A_F. To define the probabilistic rule, we will modify the policy used in Csirik et al. (2006).

The first step is to map the arrival to a bin X in the optimal fractional packing P^*. Then: Given a bin X in the optimal packing P^*, and the current arbitrary packing P, we first find an ordering $x_1, x_2, \ldots, x_{|X|}$ of the items in X, and a threshold index $last(X),\ 0 \leq last(X) \leq |X|$, such that :

- P has bins with level h_i, $0 \leq i \leq last(X)$ where:

$$h_0 = 0; \quad h_i = h_{i-1} + x_i$$

- P has no bins of levels h'_i for any $i > last(X)$ where:

$$h'_i = h_{last(X)} + x_i$$

If the arrival $j \leq last(X)$, it is placed in a bin of level h_{j-1} to create a bin of level h_j. If $j > last(X)$, it is placed in a bin of level $h_{last(X)}$.

Next we perform a case analysis to bound the change in the Lagrangian conditioned on the arrival happening from bin X of packing P^* and getting mapped to index j in the ordering:

- Case $j > last(X)$: all these items are sent to a bin of level $h_{last(X)}$ (or 0 if $last(X) = 0$). While this seemingly increases the Lagrangian since we destroy bins of some level with a higher probability than we create it, and additionally if $last(X) = 0$ then we open new bins at a probability higher than $b(F)$, we also create bin of some level which does not exist yet. This causes sufficiently large drop in potential.

—Subcase $last(X) = 0$: In this case the change in Lagrangian is upper bounded by:

$$\begin{aligned}\Delta\mathcal{L} &\leq 1 + \frac{\kappa}{\epsilon}(e^{-\epsilon} - 1)\\ &= 1 - \kappa\left(1 - \frac{\epsilon}{2!} + \frac{\epsilon^2}{3!} - \dots\right)\end{aligned}$$

for $\epsilon < 1$:

$$\begin{aligned}&\leq 1 - \kappa\left(1 - \frac{\epsilon}{2}\right)\\ &= 1 - \kappa + \kappa \cdot \frac{\epsilon}{2}\end{aligned}$$

which for $\kappa \geq 1$

$$\leq \kappa\frac{\epsilon}{2}$$

—Subcase $last(X) > 0$: In this case, we send an item to a level $h_{last(X)}$ bin with $n \doteq N_{h_{last(X)}} > 0$ so there is no change in the objective function term since no new bin is created. Further, the potential term does not increase because we create a bin of a scarcer level:

$$\begin{aligned}\Delta\mathcal{L} &\leq \frac{\kappa}{\epsilon}\left[\left(e^{-\epsilon} - 1\right) + \left(e^{-(n-1)\epsilon} - e^{n\epsilon}\right)\right]\\ &= \frac{\kappa}{\epsilon}\left(e^{\epsilon} - 1\right)\left(e^{-n\epsilon} - e^{-\epsilon}\right)\\ &\leq 0\end{aligned}$$

- Case $1 \leq j \leq last(X)$: Exactly one of these items causes a new bin to open and increase the objective function term by 1, the others cause the objective function term to change by 0. We now focus on the change in potential function due to bins of levels $h_1, \dots, h_{last(X)-1}$. We do not need to look at $h_{last(X)}$ since $j = last(X)$ only causes $N_{h_{last(X)}}$ to increase and hence potential to fall, and we have already upper bounded the effect of $j > last(X)$ on the Lagrangian.

Focusing on this set of items, the probability of increase of N_{h_i} is at least as much as the probability of decrease of N_{h_i}. Therefore, the expected change in the potential function due to level h_i is at most (let $N_{h_i} = x$ before arrival of the tagged item):

$$\begin{aligned}\Delta V_i &\leq \frac{\kappa}{2\epsilon}\left[\left(e^{-(x+1)\epsilon} - e^{-x\epsilon}\right) + \left(e^{-(x-1)\epsilon} - e^{-x\epsilon}\right)\right]\\ &= \kappa\frac{e^{-x\epsilon}}{2\epsilon}\left[e^{-\epsilon} + e^{\epsilon} - 2\right]\\ &= \kappa\frac{e^{-n\epsilon}}{\epsilon}\left[\frac{\epsilon^2}{2!} + \frac{\epsilon^4}{4!} + \frac{\epsilon^6}{6!} + \dots\right]\end{aligned}$$

for $0 < \epsilon < 1$

$$\leq \kappa\epsilon.$$

Now we look at the telescoping sum of changes in Lagrangian:

$$\Delta F_{PD}(t) + \Delta V_{PD}(t) \leq \Delta F_{A_F} + \Delta V_{A_F}. \tag{11}$$

Taking expectations:

$$\mathbf{E}[\Delta F_{PD}(t)] \leq \mathbf{E}[\Delta F_{A_F}(t)] + \mathbf{E}[\Delta V_{A_F}(t)] - \mathbf{E}[\Delta V_{PD}(t)]$$

or,

$$\mathbf{E}[\Delta F_{PD}(t)] \leq b(F) + \kappa\epsilon - \mathbf{E}[\Delta V_{PD}(t)]$$

where recall $b(F)$ is the optimal bin-rate for distribution F. Summing:

$$\mathbf{E}[F_{PD}(T)] \leq T \cdot b(F) + T\kappa\epsilon + \mathbf{E}[V_{PD}(0)] \tag{12}$$

or,

$$\mathbf{E}[F_{PD}(T)] \leq T \cdot b(F) + T\kappa\epsilon + \frac{B\kappa}{\epsilon}.$$

Setting $\kappa = 1$ and $\epsilon = \sqrt{\frac{B}{T}}$ we get:

$$\mathbf{E}[F_{PD}(T)] \leq T \cdot b(F) + 2\sqrt{BT}.$$

A.2. Proof of Theorem 2

The proof carries through similarly as in (11)-(12) with minimal changes in the last step. Changes in the potential term of the PD algorithm do not telescope any more. In particular, the potential function can decrease over time as $\epsilon(t)$ decreases.

The sum of potential differences $\sum \Delta V_{PD}(t)$ equals:

$$\begin{aligned}\sum \Delta V_{PD}(t) &= \kappa \sum_h \sum_{t=1}^{T} \frac{1}{\epsilon(t)} (e^{-\epsilon(t)N_h(t)} - e^{-\epsilon(t)N_h(t-1)}) \\ &= \kappa \sum_h \left[\frac{e^{-N_h(T)\epsilon(T)}}{\epsilon(T)} - \frac{e^{-\epsilon(0)N_h(0)}}{\epsilon(0)} + \sum_{t=1}^{T-1} \left(\frac{1}{\epsilon(t)} e^{-\epsilon(t)N_h(t)} - \frac{1}{\epsilon(t+1)} e^{-\epsilon(t+1)N_h(t)} \right) \right].\end{aligned}$$

Since the second summand in the previous expression is minimized when $n_h(T)=0$, we obtain

$$\sum \Delta V_{PD}(t) \geq \kappa \sum_h \left[-\frac{e^{-\epsilon(0)h_c(0)}}{\epsilon(0)} + \sum_{t=1}^{T-1} \left(\frac{1}{\epsilon(t)} - \frac{1}{\epsilon(t+1)} \right) \right],$$

which, after selecting $\epsilon(t) = \frac{k}{\sqrt{t+a}}$, implies

$$\sum \Delta V_{PD}(t) \geq -\frac{\kappa B}{k} \left[\sqrt{a} + \sum_{t=1}^{T-1} (\sqrt{t+1+a} - \sqrt{t+a}) \right] \geq -\kappa B \frac{\sqrt{a+T}}{k}. \tag{13}$$

Next, (12) and $\epsilon(t) = \frac{k}{\sqrt{t+a}}$ imply

$$\begin{aligned} \sum \mathbf{E}[\Delta F_{PD}(t) + \Delta V_{PD}(t) - b(F)] &\leq \sum \mathbf{E}[\Delta F_{A_F} + \Delta V_{A_F} - b(F)] \leq \kappa \sum_t \epsilon(t) = \kappa \sum_{t=1}^{T} \frac{k}{\sqrt{t+a}} \\ &\leq \kappa k \left(\frac{1}{a+1} + \int_{x=a+1}^{n=a} \frac{dx}{\sqrt{x}} \right) \\ &= \kappa k \left(\frac{1}{\sqrt{a+1}} + 2\sqrt{T+a} - 2\sqrt{a+1} \right) \leq 2\kappa k \sqrt{T+a}. \end{aligned}$$

Therefore, the previous expression and (13) give an upper bound for the expected overall additive suboptimality

$$2\kappa k \sqrt{T+a} + \kappa B \frac{\sqrt{T+a}}{k},$$

which, by setting $\kappa = 1$, $k = \sqrt{B/2}$ and $a = B$, guarantees $\epsilon(t) < 1$ and

$$\mathbf{E}[F_{PD}(T)] \leq T \cdot b(F) + \sqrt{8B\,(T+B)}.$$

A.3. Proof of Theorem 3

We will use the following version of Azuma-Hoeffding inequality.

Lemma 1. *(Azuma-Hoeffding) Suppose $X_0, \ldots, X_T$ is a supermartingale with $|X_i - X_{i-1}| \leq C_i$ for each i. Then*

$$\mathbf{Pr}[X_T - X_0 \geq a] \leq e^{-\frac{a^2}{2\sum C_i^2}}$$

Define

$$X_t = \sum_h N_h(t) + \frac{\kappa}{\epsilon} \sum_h e^{-\epsilon N_h} - \left(t \cdot b(F) + \sqrt{Bt} + \sqrt{BT} \right)$$

Then $\{X_t\}$ is a supermartingale with bounded differences:

$$\begin{aligned} |X_t - X_{t-1}| &\leq 1 + b(F) + \sqrt{B} \cdot \frac{1}{2\sqrt{t}} + 2\kappa \\ &\leq 1 + b(F) + \frac{1}{2}\sqrt{B} + 2\kappa \doteq \delta \end{aligned}$$

Which gives:

$$\mathbf{Pr}\left[\mathcal{L}(t) - \left(t \cdot b(F) + \sqrt{Bt} + \sqrt{BT} \right) \geq a \right] \leq e^{-\frac{a^2}{2t\delta}}$$

This in turn implies

$$\mathbf{Pr}\left[\sum_h N_h(t) \geq tb(F) + \sqrt{Bt} + \sqrt{BT} + \sqrt{2\lambda\delta t \log t} \right] \leq \frac{1}{t^\lambda}.$$

A.4. Proof of Theorem 4

The proof is a very minor modification over Theorem 1. Note that the crux of the argument is to show that the expected one step change in the Lagrangian is $b(F)+\epsilon$.

We will modify the policy A_F used in the proof of Theorem 1 to use the inventory state $\widetilde{N}_h(t)$ instead of the true state $N_h(t)$. The three main issues we need to iron are:

1. The expected change in potential when $\widetilde{N}_h(t)$ increases by ± 1 with equal probability: Elementary calculations show this to be ϵ.

2. If $\widetilde{N}_h(t)=\eta$, then $\tilde{N}_h(t+1)$ is either equal to $\widetilde{N}_h(t)-1$, or $\widetilde{N}_h(t)$: This causes only an increase in the potential term. However, in this situation, the change is bounded by $\frac{\epsilon}{2}\left((\eta-\eta+1)^2-(\eta-\eta)^2\right)=\frac{\epsilon}{2}$.

3. When $last(X)=0$, then a new bin is opened with probability 1: In this case we want the decrease in potential to annihilate the increase of 1 in $\sum_h N_h(t)$. Since all the bins created in this case have $\widetilde{N}_h(t)=0$, the total decrease in potential is: $\frac{\epsilon}{2}\left[(\eta-1)^2-(\eta-0)^2\right]=-\frac{\epsilon(2\eta-1)}{2}\leq -\epsilon(\eta-1)\leq -1$ for $\eta\geq\frac{1}{\epsilon}+1$.

A.5. Proof of Theorem 5

Since the item size distribution has a lot of structure, we can arrive at the optimal online algorithm when there is a bound I on the number of open bins. The first observation is that if there is a level 2 bin and a size 1 item arrives then the item must be placed in the level 2 bin. Similarly, if a size 2 item arrives and there is a level 1 bin then it must be placed in the level 1 bin. Therefore the only case in which the inventory of open bins can have both a level 1 and a level 2 bin is when an item of size 1 arrives to find all bins in inventory of level 1. Subsequently, the number of level 2 bins can at most be 1 if there are any level 1 bins in inventory. We can thus describe the state of inventory as a Markov chain with state $N(t)=(N_1(t),N_2(t))$, where $N_h(t)$ is the number of level h bins in inventory at time t. The possible state transitions are given by (all with probability 1/2 each):

$$\begin{aligned}
&(0,0)\xrightarrow{1}(1,0)\ ;\ (0,0)\xrightarrow{2}(0,1)\\
a<I:\ &(a,0)\xrightarrow{1}(a+1,0)\ ;\ (a,0)\xrightarrow{2}(a-1,0)\\
b<I:\ &(0,b)\xrightarrow{1}(0,b-1)\ ;\ (0,b)\xrightarrow{2}(0,b+1)\\
&(I,0)\xrightarrow{1}(I-1,1)\ ;\ (I,0)\xrightarrow{2}(I-1,0)\\
&(0,I)\xrightarrow{1}(0,I-1)\ ;\ (0,I)\xrightarrow{2}(0,I)\\
&(a,1)\xrightarrow{1}(a,0)\ ;\ (a,1)\xrightarrow{2}(a-1,1)
\end{aligned}$$

It is now easy to argue that for $t=\omega(I^2)$, $\mathbf{Pr}[X(t)=(0,I)]\geq\frac{1}{3I}+o(1)$, and therefore, the probability that an arriving size 2 item sees all bins in inventory of level 2 and is placed in a bin by itself is at least $\delta=\frac{1}{3I}$. Therefore, the expected number of bins used will be at least $\frac{T}{2}(1-\delta)+\frac{T}{2}\delta+\frac{T}{2}\delta\cdot\frac{1}{3}=\frac{T}{2}\left(1+\frac{\delta}{3}\right)$. Therefore, to get $(1+\epsilon)$ optimal solution, $I\geq\frac{1}{9\epsilon}$.

A.6. Proof of Theorem 6

Let $S_1, S_2, \ldots, S_T$ be independent random variables samples according to distributions $q_1, \ldots, q_T$. Let $\Delta\mathcal{L}(t) = \mathcal{L}(t) - \mathcal{L}(t-1)$. Our goal is to bound

$$\begin{aligned}\mathbf{E}[\mathcal{L}(T) - \mathcal{L}(0)|\mathcal{F}_t] &= \mathbf{E}\left[\sum_{t=1}^{T} \Delta\mathcal{L}(t)|\mathcal{F}_t\right] \\ &= \mathbf{E}\left[\sum_{t=1}^{T} \Delta\mathcal{L}(t) \sum_{k=1}^{t} q_k(t)|\mathcal{F}_t\right] \quad = \mathbf{E}\left[\sum_{k=1}^{T}\sum_{t=1}^{T} \Delta\mathcal{L}(t) q_k(t)|\mathcal{F}_t\right] = \sum_k \mathbf{E}[\Delta\mathcal{L}(S_k)|\mathcal{F}_t]\end{aligned}$$

We will focus on the analyzing the $\mathbf{E}[\Delta\mathcal{L}(S_k)|\mathcal{F}_t]$ for a given random time S_k, and let q_k denote the corresponding measure. By definition of S_k, the arriving job at time S_k is a random sample from the smoothed distribution $\hat{F}_{\pi,k|t}$. However we can not use the A_F policy described earlier corresponding to $\hat{F}_{\pi,k|t}$ since we had assumed that the packing seen by size j item is independent of j. This is not true anymore.

Instead, we will exploit the fact that support of S_k is on an interval $[\ell_k, \ell_k+1, \ldots, \ell_k+L-1]$. Therefore, while the packing seen by an arrival at time S_k is random and possibly correlated with the item size, all packing states in the support of this distribution differ by at most L arrivals, and therefore are close to each other.

We now formalize the argument. Denote the state of the packing just before time t by P_t. The preceding argument implies

$$\forall k: \sup_{t_1, t_2 \in [\ell_k, \ldots, \ell_k+L-1]} ||P_{t_1} - P_{t_2}|| \leq L \tag{14}$$

where $||P_{t_1} - P_{t_2}|| \doteq \max_{1\leq h\leq B} |N_h(t_1) - N_h(t_2)|$. That is, for a fixed k, the possible set of packings that an arrival at random time S_k sees is supported on an L-narrow set. We make no further assumption on the joint distribution of the size of the arriving item and the packing seen by the arrival.

We now invoke Lemma 2:

$$\mathbf{E}_{S_k}[\Delta\mathcal{L}(S_k)] \leq \hat{b}_{\pi,k|t} + \kappa\epsilon\left(\frac{2L+3}{4}\right). \tag{15}$$

A telescoping sum argument as before gives the bound in the theorem statement.

LEMMA 2. *Consider the following one-step bin packing problem: An item type is sampled from distribution F, with $b(F)$ denoting the optimal bin-rate under F returned by the packing LP. If the item is of type j, an initial packing is sampled according to a measure μ_j on a set of packings $\mathcal{U}$. The universe of initial packings $\mathcal{U}$ is the same for all types j, and further satisfies the condition that it is* L-narrow*, meaning:*

$$\forall\ P_1, P_2 \in \mathcal{U}\ :\ \max_h |N_h(P_1) - N_h(P_2)| \leq L$$

On seeing the item type $Y \sim F$, and the initial packing $P \sim \mu_Y$, the item is packed using the Primal-Dual algorithm with Lagrangian function:

$$\mathcal{L}(\mathbf{N}) = \sum_h N_h + \frac{\kappa}{\epsilon} \sum_h e^{-\epsilon N_h} \tag{16}$$

where $L < \frac{1}{2\epsilon} \log \frac{\kappa B}{B-1}$ (equivalently $\kappa e^{-2\epsilon L} \geq 1 - \frac{1}{B}$) and $\epsilon L < 1$, to obtain the final packing P'. Then:

$$\mathbf{E}_{Y \sim F} \left[\mathbf{E}_{P \sim \mu_Y} \left[\mathcal{L}(P') - \mathcal{L}(P) \right] \right] \leq b(F) + \kappa \epsilon \left(\frac{2L+3}{4} \right). \tag{17}$$

Proof: Following the outline of the previous results, we will rely on the fact that the Primal-Dual algorithm minimizes the Left Hand Side of (17), and bound it by the drift of a randomized policy. We will pick a randomized policy very similar to the A_F policy that (i) starts with an optimal packing P^* corresponding to F, (ii) maps the incoming arrival to a random item in P^* (say in bin X), and (iii) then proceeds to pack this item in P accordingly. However, in the prior, the current packing P was deterministic, and this was used in step (iii) because each action of A_F was a feasible action for the Primal-Dual algorithm.

Since we want the randomized policy A_F to only take actions which are almost surely feasible for Primal-Dual, we define the following *meet* packing corresponding to $\mathcal{U}$ (the term meet comes from orders on lattices, where meet denotes the greatest lower bound of a partially ordered set of elements):

$$\underline{P} \doteq \bigwedge \mathcal{U} \doteq P_1 \wedge P_2 \wedge \cdots \wedge P_{|\mathcal{U}|} \tag{18}$$

That is:

$$N_h(\underline{P}) = \min_{P \in \mathcal{U}} N_h(P)$$

The A_F policy is now almost the same as before, except we use the packing $\underline{P}$ to decide the placement of the item, and require a stricter lower bound on when we consider a level to be non-existent. Two crucial observations are:

1. $N_h(P) \geq N_h(\underline{P})$ (almost surely): Therefore if it is feasible for A_F to send an item to level h, then it is also feasible with probability 1 for our Primal-Dual algorithm to do so.

2. $N_h(P) \leq N_h(\underline{P}) + L$ (almost surely): For the packing $\underline{P}$, we expect the drift in the Lagrangian to be $b(F) + \Theta(\epsilon)$. Since the initial packing is L-close to $\underline{P}$, we will prove that the drift only worsens by a small amount.

We will use the following tweaked policy A_F:

Given a bin X in the optimal packing P^* (that the arriving size maps to) and the packing $\underline{P}$, we first find an ordering $x_1, x_2, \ldots, x_{|X|}$ of the items in X, and a threshold index $last(X)$, $0 \leq last(X) \leq |X|$, such that if we set $h_i \equiv \sum_{j=1}^{i} x_j$, then:

- $\underline{P}$ has at least L partially filled bins with each level h_i, $0 \le i \le last(X)$
- $\underline{P}$ has at most $L-1$ partially filled bins in level $h_{last(X)} + x_i)$ for any $i > last(X)$

Below we show the change in Lagrangian for various cases, conditioned on the arrival happening from bin X of optimal packing and getting mapped to index j in the ordering:

- Case $last(X) = 0$: In this case all items from X open a new bin, but in $\underline{P}$ there are at most $L-1$ bins of the levels created thus, and therefore in the true P there may be up to $2L$ bins of the new level created. Therefore the change in Lagrangian is:

$$\begin{aligned}\Delta\mathcal{L} &\le 1 + \frac{\kappa}{\epsilon} e^{-\epsilon 2L}(e^{-\epsilon} - 1) \\ &= 1 - \kappa e^{-\epsilon 2L}\left(1 - \frac{\epsilon}{2!} + \frac{\epsilon^2}{3!} - \dots\right)\end{aligned}$$

for $\epsilon < 1$:

$$\le 1 - \kappa e^{-2\epsilon L}\left(1 - \frac{\epsilon}{2}\right)$$

since $\kappa e^{-2\epsilon L} \ge 1 - \frac{1}{B}$

$$\begin{aligned}&\le \frac{1}{B} + \kappa\frac{\epsilon}{2} \\ &\le b(F) + \kappa\frac{\epsilon}{2}\end{aligned}$$

- Case $last(X) > 0$

—Subcase $1 \le j \le last(X)$: Exactly one of these items causes a new bin to open and increase the objective function term by 1, the others cause the objective function term to change by 0. We now focus on the change in potential function due to bins of levels $h_1, \dots, h_{last(X)-1}$. We do not need to look at $h_{last(X)}$ since $j = last(X)$ only causes this number to increase and hence potential to fall, and we will account for the effect of $j > last(X)$ on the Lagrangian in the next subcase.

Focusing on this set of items, the probability of the increase of N_{h_i} is at least as much as the probability of decrease of N_{h_i}. However, the actual value of N_{h_i} when it increases by 1 and when it decreases by 1 can be different by up to L (example, N_{h_i} decreases on arrival of item of size 2 for which case $N_{h_i} = L$, and increases on arrival of size 3 for which case $N_{h_i} = 2L+1$).
It is easy to see that the above case is indeed the worst for the change in the Lagrangian, and therefore the expected change due to level h_i is almost surely bounded by :

$$\begin{aligned}\Delta V_i &\le \frac{\kappa}{2\epsilon}\left[\left(e^{-(2L+1)\epsilon} - e^{-2L\epsilon}\right) + \left(e^{-\epsilon(L-1)} - e^{-\epsilon L}\right)\right] \\ &= \frac{\kappa e^{-\epsilon L}}{2\epsilon}\left[e^{-\epsilon L}(e^{-\epsilon} - 1) + (e^{\epsilon} - 1)\right]\end{aligned}$$

Since $\epsilon < 1$ and $\epsilon L < 1$

$$
\begin{aligned}
&\leq \frac{\kappa e^{-\epsilon L}}{2\epsilon}\left[(1-\epsilon L)\left(-\epsilon+\frac{\epsilon^2}{2}\right)+(\epsilon+\epsilon^2)\right] \\
&\leq \frac{\kappa}{2}\left[(1-\epsilon L)(-1+\frac{\epsilon}{2})+1+\epsilon\right] \\
&= \frac{\kappa}{2}\epsilon\left(L+\frac{3}{2}\right)
\end{aligned}
$$

—Subcase $j > last(X)$: all these items are sent to a bin of level $h_{last(X)}$. In the process, we kill bins of level $h_{last(X)}$ (note that we must have at least L bins of this level) with a higher probability than we create it and that increases our Lagrangian. However, similar to the case $last(X) = 0$, we also create bins of levels for which $\underline{P}$ has at most $L-1$ bins, and therefore, in the true P, we have at most $2L-1$ bins. There is no change in the objective function term, and a similar analysis as the previous Subcase shows that the potential term increases by at most:

$$\Delta\mathcal{L} \leq \frac{\kappa}{2}\epsilon\left(L+\frac{3}{2}\right)$$

□