

Online Demand Scheduling with Failovers

Konstantina Mellou
Microsoft Research
kmellou@microsoft.com

Marco Molinaro
Microsoft Research
molinaro.marco@gmail.com

Rudy Zhou *
Carnegie Mellon
rbz@andrew.cmu.edu

Abstract

Motivated by cloud computing applications, we study the problem of how to optimally deploy new hardware subject to both power and robustness constraints. To model the situation observed in large-scale data centers, we introduce the *Online Demand Scheduling with Failover* problem. There are m identical devices with capacity constraints. Demands come one-by-one and, to be robust against a device failure, need to be assigned to a pair of devices. When a device fails (in a failover scenario), each demand assigned to it is rerouted to its paired device (which may now run at increased capacity). The goal is to assign demands to the devices to maximize the total utilization subject to both the normal capacity constraints as well as these novel failover constraints. These latter constraints introduce new decision tradeoffs not present in classic assignment problems such as the Multiple Knapsack problem and AdWords.

In the worst-case model, we design a deterministic $\approx \frac{1}{2}$ -competitive algorithm, and show this is essentially tight. To circumvent this constant-factor loss, which in the context of big cloud providers represents substantial capital losses, we consider the stochastic arrival model, where all demands come i.i.d. from an unknown distribution. In this model we design an algorithm that achieves a sub-linear additive regret (i.e. as OPT or m increases, the multiplicative competitive ratio goes to 1). This requires a combination of different techniques, including a configuration LP with a non-trivial post-processing step and an online monotone matching procedure introduced by Rhee and Talagrand.

*Work done while visiting Microsoft Research Redmond

Contents

1	Introduction	1
1.1	Our results	3
1.2	Technical Overview	4
1.3	Related work	5
2	FAILOVER Problem in the Online Worst-Case Model	6
2.1	Algorithm	6
2.2	Analysis	7
3	Sublinear Additive Regret in the Stochastic Model	9
3.1	Algorithm	10
3.2	Analysis	11
4	Offline Machine Minimization	16
4.1	Configuration LP	16
4.2	Matching configurations	17
4.3	Reducing the number of types and demands in a configuration	20
4.4	Putting it all together	22
5	Rate of Convergence of the Minimum Number of Machines	24
5.1	Proof of Lemma 8 : $\text{OPT}_{mach}(\mu_T)$ is a good proxy	25
5.2	Proof of Lemma 9 : Approximate linearity of $\text{OPT}_{mach}(\mu_T)$	26
A	Concentration Inequalities	31
B	Auxiliary Results	31
C	Additional Results for Online Worst-Case	32
C.1	Upper bound	32
C.2	Algorithm for small demands	33
D	Omitted Proofs from Section 3	35
D.1	Proof of Lemma 1	35
D.2	Proof of Lemma 3	36
E	Properties of (LP_{mach})	38
E.1	Consistency of LP_{mach} with respect to type partitions	38
E.2	Solving (LP_{mach})	39

1 Introduction

A critical challenge faced by cloud providers is how to optimally deploy new hardware to satisfy the ever increasing demand for cloud resources, and the main bottleneck in this process is power. Data centers consist of power devices with limited capacity and each demand for hardware (e.g., rack of servers) has a power requirement. The goal is to assign demands to power devices to fulfill their requirements while using the available power in the data centers efficiently. This allows cloud providers to maximize their return on investment on existing data centers before needing to incur large capital expenses for new data centers to accommodate additional demand.

An important consideration that sets this demand assignment process apart from other applications is *reliability*. Cloud users are promised a high availability of service which mandates that cloud capacity can only be unavailable for very short durations (between a few minutes and a few hours per year). As a result, assigning each demand to a single power device leads to an unacceptable level of risk; if that device fails, the capacity for the demand becomes unavailable, leading to potentially millions of dollars in costs for the provider and jeopardizing the cloud business model that is highly dependent on users' trust. To this end, power *redundancy* is built into the assignment process.

We consider a specific model of redundancy used by large cloud providers [22]. In this model, each demand gets assigned to two power devices. In normal operations (no device failure), the demand obtains half of its required power from each device. If one of the devices fails, the remaining device must provide the full power amount to the demand (see Figure 1 for an example). In these failover scenarios, the remaining devices may run at an increased capacity temporarily to accommodate their increased load. The provider uses this time to take ad-hoc corrective actions, for instance, shut down certain workloads and reduce the power of others in order to bring the power utilization of each device back within its normal limits. As in [22] we consider a single device failure at a time.

This architecture is favored in practice because it provides strong reliability guarantees with a small increase in overhead and complexity. One could consider more complex architectures, where demands could be assigned with a power split other than half-half to each device or to more than two devices, but this comes at an increased cost in hardware and operational complexity. Further, a common goal of large cloud providers is to provide statistical guarantees for high service availability, e.g., 99.99% availability for certain cloud resources or services; cloud operators have determined that accounting for a single device failure with the described architecture provides such target guarantees.

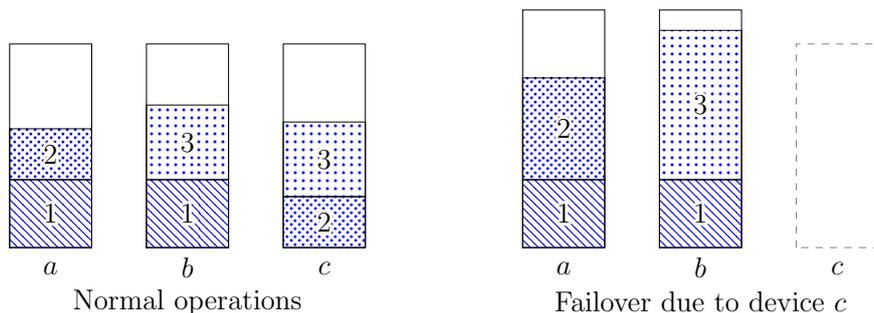


Figure 1: In normal operations (left), each demand (denoted with a different pattern) is assigned to two devices and gets half of its required power from each device. In the failover scenario where device c has failed (right), the demands that were assigned to c now get their full power from the remaining devices that may run at increased capacity.

We introduce the *Online Demand Scheduling with Failover* problem (FAILOVER) to model this issue of assigning demands to power devices with redundancy. Formally, in this problem there are m identical devices (or machines) and n demands. Each device has two capacities: a nominal capacity that is normalized to 1 and a failover capacity $B \geq 1$. Each demand j has some size $s_j \geq 0$, which for convenience is defined as its per-device power requirement (so the total power requirement of the demand is $2s_j$). The demands arrive online one-by-one and there is no knowledge about future demands. The goal is to irrevocably assign the arriving demands to pairs of devices (or edges, where we consider each device as a node) satisfying:

1. (Nominal Constraints) For every device u , its total load has to be at most 1, namely $L_u := \sum_{v \neq u} L_{uv} \leq 1$, where we define $L_{uv} = \sum_{j \rightarrow uv} s_j$ to be the total load on edge uv (i.e., all demands assigned to the pair of devices uv).
2. (Failover Constraints) For every device u , we have $L_u + \max_{v \neq u} L_{uv} \leq B$ (i.e., if a device $v \neq u$ fails, all demands assigned to uv have to be supplied solely by u , which sees its load increased by the amount L_{uv} that was formerly supplied to them by device v ; the increased load has to fit the failover capacity B).

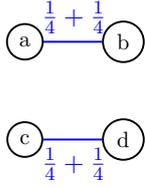
We assume that each demand size s_j fits on a pair of devices by itself, so $s_j \in [0, \min(1, B/2)]$. We are not allowed to reject demands, so the algorithm assigns arriving demands to the available devices until a demand cannot be scheduled, in which case the algorithm terminates. Our objective is to maximize the total size of all assigned demands (i.e., the utilization). We compare the algorithm against the optimal offline strategy that knows the demand sequence in advance (but still subject to the same no-rejection requirement). We use OPT to denote the total utilization of this optimal offline strategy.

This problem has similarities with several classical packing problems. For example, in the Multiple Knapsack problem (and related problems such as Generalized Assignment [21], AdWords [15], etc.) we are given a set of items each with a weight and size, and the goal is to select a subset of the items to pack in capacitated bins in order to maximize the total weight. However, one fundamental difference in our setting, besides the need to assign each demand to a pair of devices instead of a single device, is the failover constraint. Unlike in previously studied resource allocation problems, here the capacity constraints are not just determined by the total demand incident to a node, but rather they depend also on how the demands are arranged across its edges. See the next example.

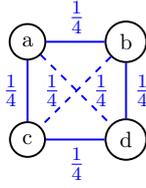
Example 1. Consider an instance with 4 power devices a, b, c, d with failover capacity $B = 1$. There are 6 demands of size $\frac{1}{4}$ that arrive sequentially; suppose 4 demands have arrived so far. One possible assignment has placed 2 demands on each of the pairs ab and cd (see Figure 2a), while a different one may place each of the 4 demands on a different pair (see Figure 2b, in solid lines). While in the second option all remaining demands can be placed (dashed lines in Figure 2b), the first option cannot accommodate more demands due to the Failover capacity. To see this, suppose we assign another demand to device a , say. If device b fails, then the total load on a will become at least $\frac{5}{4}$ violating its Failover capacity.

The above example suggests that due to the Failover constraints we should “spread out” the demands by not putting too many demands on one edge, because if one of its endpoints fails then this edge can have a large contribution to the Failover constraint of the other endpoint. However, there is a danger in spreading out the demands too much and not leaving enough devices free.

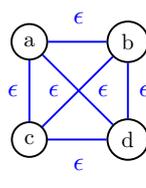
Example 2. Consider again the same 4 power devices a, b, c, d with failover capacity $B = 1$. Now, there are 7 demands; the first 6 have a small size $\varepsilon > 0$ and the last demand has size 0.5.



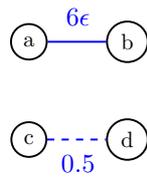
(a) Bad assignment.



(b) Good assignment.



(a) Bad assignment.



(b) Good assignment.

Figure 2: Illustration of [Example 1](#).Figure 3: Illustration of [Example 2](#).

Assume the first 6 demands have arrived. A first option is to assign one demand of size ϵ per device pair (see [Figure 3a](#)). In this case, the remaining demand of size 0.5 cannot be placed, as the Failover capacities would be exceeded. The second option groups the first 6 demands on a single edge (see [Figure 3b](#)); in this case, all demands can be fulfilled by assigning the last demand on a disjoint edge (dashed edge of [Figure 3b](#)).

Taking these two examples together, we see that there is a delicate balance between spreading demands out across edges to minimize their impact in failover scenarios and leaving enough devices open for future demands, as to not prematurely end up with an unassignable demand.

1.1 Our results

We start by considering the FAILOVER problem in the worst-case and design a deterministic algorithm with competitive ratio $\approx \frac{1}{2}$. Since no deterministic algorithm can be better than $\frac{1}{2}$ -competitive ([Theorem 10](#) in [Appendix C.1](#)), this result is almost best possible. (For the special case where demand sizes are small, we adapt our algorithm to obtain an improved competitive ratio, see [Theorem 11](#).)

Theorem 1. *There is a deterministic poly-time online algorithm for FAILOVER in the worst-case model with competitive ratio $\frac{1}{2} - O(\frac{1}{m^{1/3}})$,¹ where m is the number of devices.*

A $\frac{1}{2}$ -competitive solution may, roughly speaking, underutilize by a factor of $\frac{1}{2}$ the available power; in the context of big cloud providers, this inefficiency translates to substantial capital expenses due to the extra data centers required to accommodate the demands. Since such losses are unavoidable in the worst-case model, we consider the FAILOVER problem in the *stochastic arrival model*. Here the demand sizes are drawn i.i.d. from an unknown distribution μ supported on $[0, \min(1, B/2)]$.

We show that in this stochastic model it is possible to obtain *sublinear additive regret*. This means that as OPT (or, equivalently, the number of devices) grows, the multiplicative competitive ratio of our algorithm goes to 1.

Theorem 2. *For the FAILOVER problem in the stochastic arrival model, there is a poly-time algorithm that achieves utilization $\text{OPT} - O(\text{OPT}^{5/6} \log \text{OPT})$ with probability $1 - O(\frac{1}{m})$.*

As a subroutine of this algorithm, we need to solve the natural *offline minimization* variant of demand scheduling with failover: Given a collection of demands, minimize the number of devices needed to assign all demands satisfying the Nominal and Failover constraints. We also design an (offline) algorithm with sublinear additive regret for this problem ([Section 4](#)).

¹Throughout the paper we use $O(x)$ to mean “ $\leq cst \cdot x$ ” for some constant cst independent of x .

1.2 Technical Overview

We illustrate the main technical challenges in the FAILOVER problem in both the worst-case and stochastic models, as well as in the offline minimization subproblem needed for the latter.

Online Worst-Case (Section 2). The examples from Figure 2 and 3 show that the main difficulty is dealing with the trade-off between spreading out the demands, which allows for a better use of the failover budgets, and co-locating demands on fewer edges, keeping some edges free for future big demands.

To effectively strike this balance and get near optimal guarantees, the main idea is to group demands based on their sizes using intervals I_k and schedule each group separately on cliques of size k . That is, we will “open” a set of k unused devices and assign the demands in I_k only to the edges between these devices (opening new k -cliques as needed). Interestingly, we assign at most one demand per edge of the clique (other than for tiny demands, which are handled separately). This means the algorithm tries to co-locate demands in controlled *regions*, which allows for the right use of the failover budgets.

Online Stochastic Arrivals (Section 3). First, note that because demands are i.i.d. from a distribution with bounded support, the total utilization of the first ℓ demands grows as $\ell \cdot \mathbb{E}_{S \sim \mu} S$. Thus, it suffices to show that our algorithm “survives” for as many demand arrivals as possible without needing to reject one due to lack of space. Our approach is to try and assign prefixes of arrivals to the (approximately) minimum number of devices possible. This ensures that if our algorithm fails due to needing more than m devices to feasibly assign another demand, then OPT will fail shortly after.

Our algorithm is based on a learn-and-pack framework, where we use knowledge of the first ℓ arrivals to compute a good template assignment for the next ℓ arrivals. To compute this template, we need a subroutine that (approximately) solves the offline minimization subproblem mentioned above. Concretely, we run the subroutine on the realized sizes of the first ℓ arrivals, which gives a possible assignment of these demands into, say m' unused devices. We use the “slots” of this possible assignment as a template to assign the future ℓ demands by employing the *online monotone matching* process of Rhee-Talgrand [19]: For each future arrival, we assign it to a (carefully-chosen) open slot in the template that has a larger size – if we cannot find such an open slot, then we assign this demand to its own disjoint edge (using 2 more devices).

It is known that this matching process leaves $o(\ell)$ unmatched demands with high probability. Further, our offline minimization subroutine has sublinear additive regret, that is, it uses only $o(\ell)$ more devices than the optimal offline assignment. Since these losses are sublinear in the prefix size, it seems that by repeating this process together with doubling the prefix size we should obtain a final sublinear regret guarantee.

But there is still a major issue: This strategy uses *disjoint* sets of devices to fulfill the first ℓ demands and the next ℓ demands (for each doubling ℓ). But this is possibly very wasteful: even using the optimal assignment for each of these ℓ demands separately may require many more devices (up to double) compared to reusing the leftover space from the first batch of ℓ demands for the next batch (i.e. assigning the batches to a common set of devices). Wasting a constant fraction of devices would lead to the unwanted constant-competitive loss. To overcome this, we show that M_ℓ , the minimum number of devices to assign ℓ i.i.d. demands, is approximately linear in ℓ (Section 5), e.g. $M_\ell + M_\ell$ (assigning batches separately) is approximately $M_{2\ell}$ (assigning them together). This is a non-trivial task (another Rhee-Talagrand paper [18] is entirely devoted to doing this for the simpler Bin Packing problem). Perhaps surprisingly, our proof relies on our algorithm for the offline device minimization problem, which is LP-based.

The crucial property is that the optimal LP value doubles if we duplicate the items on its input, which (with additional probabilistic arguments) translates into the additivity of M_ℓ .

Offline Minimization (Section 4). Our algorithm for offline minimization of the number of devices needed to fulfill a set of demands is based on a configuration LP inspired by the classic Gomory-Gilmore LP for the Bin Packing problem. Consider a fixed assignment of demands to some number of devices. We want to interpret each device as a configuration, which captures the arrangement of demands on this device’s edges. Our LP will minimize the number of configurations needed in order to assign all demands.

There is a tension between two issues in this approach. First, the Failover constraint depends not only on the subset of demands on this device’s edges, but also how they are arranged within these edges (because the most-loaded edge contributes to the Failover constraint). This suggests that a configuration should not only specify a subset of demands, but also have enough information about the edge assignment to control the most-loaded edge. Second, each demand must be assigned to a pair of devices rather than a single device, so our configurations are not “independent” of each other. Thus, we need to “match” configurations to ensure that a collection of configurations can be realized in an edge assignment. In summary, our configurations should be expressive enough to capture the Failover constraints, but also simple enough so that we can actually realize them in an actual assignment.

Our solution to this is to define a configuration to be a subcollection, say C , of demands satisfying $\sum_{s \in C} s \leq 1$ (the Nominal constraint) and $\sum_{s \in C} s + \max_{s \in C} s \leq B$ (a relaxed Failover constraint). Note that this notion of configuration does not capture the arrangement of the demands C across a device’s edges – we assume the best case that every demand is on its own edge to minimize their impact in failover scenarios. It is not clear that there even exists a near-optimal assignment that assigns at most one demand per edge, let alone that we can obtain one from the LP solution. However, our LP post-processing procedure will show that – by opening slightly more devices – we can match configurations of this form to realize them in a near-optimal assignment.

1.3 Related work

Despite a vast literature on assignment-type problems, none of the ones considered addresses the main issue of redundancy, modeled in the FAILOVER problem. Arguably the Coupled Placement [11] problem is the closest to FAILOVER. Given a bipartite graph with capacities at the nodes and a set of jobs, the goal is to assign a subset of the jobs to the edges of the graph to maximize the total value (each assigned job gives a value that also depends on its assigned edge), while respecting the capacity of the nodes (each assigned job consumes capacity from its edge’s endpoints). [11] gives a $\frac{1}{15}$ -approximation to the offline version of this problem (see also [1]). While this problem involves the allocation of jobs to a pair of nodes (albeit on a bipartite graph) and has the additional difficulty that the value and consumption of a job depends on which pair of nodes it is assigned, it does not have any Failover type constraints, a crucial component of our problem.

As already mentioned, several classic assignment problems are related to ours, such as the Multiple Knapsack [4], Generalized Assignment (GAP) [21], and AdWords problem [15, 7]. The latter is the closest to our problem: there are m bins (i.e. advertisers) of different capacities, and jobs (i.e. keyword searches) that come one-by-one and need to be assigned to the bins; each assignment consumes some of the bin’s capacity and incurs an equal amount of value (i.e. bid). The goal is maximize total value subject to bin capacities. Despite the similarities, this problem does not consider critical aspects of our problem, namely the need to assign a

job/demand to a pair of bins/devices and the Failover constraints.

There is also a large literature on survivable network design problems, where failures in the network are explicitly considered [6], but the nature of the problems is quite different from our assignment problem, as the focus there is typically on routing flows.

Finally, a problem related to our device minimization problem, and from which we borrow some tools and techniques, is Bin Packing. Here jobs of different sizes need to be assigned to a minimum number of bins of size 1. Results are known in both offline [9] and online settings [2, 19]. In the online stochastic setting, [19] obtains an additive $+O(\sqrt{\text{OPT}} \cdot \log^{3/4} \text{OPT})$ sublinear approximation (see [5, 8, 13] for improvements under different assumptions).

2 FAILOVER Problem in the Online Worst-Case Model

In this section we consider the FAILOVER in the online worst-case model. We design an algorithm that achieves competitive ratio $\approx \frac{1}{2}$ in this setting (restated from the introduction).

Theorem 1. *There is a deterministic poly-time online algorithm for FAILOVER in the worst-case model with competitive ratio $\frac{1}{2} - O(\frac{1}{m^{1/3}})$,² where m is the number of devices.*

Recall that in Appendix C we also show the almost matching upper bound of $\frac{1}{2}$ on such competitive ratio, and design another algorithm whose competitive ratio approaches 1 as the size of the largest demand goes to 0. To convey the main ideas more clearly, here we focus only on Theorem 1.

2.1 Algorithm

As suggested in the technical overview, our algorithm will group demands by size, and assign each group of demands to sub-cliques of an appropriate size.

To make this precise, set in hindsight $L := m^{1/3}$ and for $k = 2, \dots, L - 1$ define the interval

$$I_k := \left(\min \left\{ \frac{1}{k}, \frac{B}{k+1} \right\}, \min \left\{ \frac{1}{k-1}, \frac{B}{k} \right\} \right].$$

(Notice there is no $k = 1$, because the upper limit of I_2 is the max size of a demand.) This definition ensures that it is feasible to assign one demand of such size to each edge of a k -clique, as we argue in the next subsection. Also define the interval of small sizes

$$I_{\geq L} := \left[0, \min \left\{ \frac{1}{L-1}, \frac{B}{L} \right\} \right].$$

The algorithm is then the following:

²Throughout the paper we use $O(x)$ to mean " $\leq cst \cdot x$ " for some constant cst independent of x .

Failover Worst Case:

- 1: When a demand arrives, determine the interval I_k (or $I_{\geq L}$) that it belongs to based on its size.
 - 2: If it belongs to an interval I_k with $k \in \{2, \dots, L-1\}$, assign the demand to any “empty” edge (i.e. that has not received any demands) of a k -clique opened for I_k . If no such edge exists, then open a new k -clique for I_k .
 - 3: Otherwise it belongs to $I_{\geq L}$, so assign it to an edge of one of its L -cliques using *first-fit* (so here we **can assign multiple demands to the same edge**) making sure that the total load on each edge is at most $\min\{\frac{1}{L-1}, \frac{B}{L}\}$. By first-fit we mean that the edges of the $I_{\geq L}$ cliques are arbitrarily ordered and the demand is assigned to the first possible edge. Open a new L -clique for $I_{\geq L}$ if need be.
 - 4: If the demand cannot fit in the appropriate clique and it is not possible to open a new clique (i.e. there are not enough unused machines to form a clique of the desired size), then stop.
-

2.2 Analysis

We first quickly verify that the assignment done by the algorithm is feasible, i.e. satisfies the Nominal and Failover constraints. Consider a node/machine u on an I_k clique opened by the algorithm (for machines in an $I_{\geq L}$ clique the argument is the analogous). For the Nominal capacity constraint: Every demand assigned to u is actually assigned to one of the $k-1$ edges in this clique incident on u ; each such edge receives at most 1 demand from I_k (and no other demands), so using the upper limit of this interval we see that u receives total size at most $(k-1) \cdot \min\{\frac{1}{k-1}, \frac{B}{k}\} \leq 1$, so within its Nominal capacity. For the Failover capacity: in a failover scenario one of these $(k-1)$ demands has “both ends” assigned to u , so the total size it receives is now $k \cdot \min\{\frac{1}{k-1}, \frac{B}{k}\} \leq B$, so within the Failover capacity. Hence the algorithm produces a feasible assignment.

Now we show that the value obtained by the algorithm is at least $(\frac{1}{2} - O(\frac{1}{m^{1/3}}))\text{OPT}$. The idea is to show that for (essentially) each clique opened by the algorithm, we get on average value at least $\approx \frac{1}{4}$ per vertex. Given that each node has Nominal capacity 1 and each demand must be scheduled on two nodes, OPT can only get at most $\frac{1}{2}$ value from each node on average, so this shows that our algorithm is a $\approx \frac{1}{2}$ -approximation. However, there are two exceptions where we may get less than $\approx \frac{1}{4}$ per vertex on average. The first is the last clique for each I_k , which may not be “fully used” (but by setting L appropriately there are not too many nodes involved in this loss). More importantly, the second exception is the “big items” I_2 , which may not allow us to get average value $\frac{1}{4}$ per node (e.g. when the failover is $B=1$, a demand of size $\frac{1}{3} + \varepsilon$ falls in the group I_2 and is put by itself on an edge, giving value $\frac{1}{6} + \frac{\varepsilon}{2} \ll \frac{1}{4}$ per node used). However, in this case we show that we can obtain a stronger upper bound for these demands for OPT .

We now make this precise. Assume throughout that the algorithm has stopped before the end of the input (else it scheduled everything, so it is OPT). We account for the value obtained on each type of clique separately.

Cliques for $I_{\geq L}$. We will use two observations:

- When the algorithm opens a new $I_{\geq L}$ clique, every edge of the previous $I_{\geq L}$ cliques has some demand assigned to it.
- Across all $I_{\geq L}$ cliques, out of all edges with some demand assigned to them, at most one can have total size assigned to it less than $\alpha := \frac{1}{2} \min\{\frac{1}{L-1}, \frac{B}{L}\}$ (i.e. half of its “capacity”).

Both observations stem from the first-fit strategy to assign these demands. In particular, the algorithm will only open a new clique when a demand in $I_{\geq L}$ does not fit in the edges of the existing cliques, implying that all of these edges already have some demand assigned; this shows the first statement. For the second statement, by contradiction assume that at some point there are at least two edges on $I_{\geq L}$ cliques with total load less than α . Then the first demand that was assigned to the last such edge has size less than α . But this means that it could have been assigned to an earlier edge with load less than α , contradicting the first-fit procedure.

Let $c_{\geq L}$ be the total number of $I_{\geq L}$ cliques that the algorithm opened, and $m_{\geq L} := c_{\geq L} \cdot L$ the number of nodes/machines associated with those cliques. Combining the above two observations, at the end of the execution either: (i) every edge of the first $c_{\geq L} - 1$ of these cliques has load at least α or; (ii) all but one edge in the first $c_{\geq L} - 1$ cliques has load at least α and some edge of the last $c_{\geq L}$ -th (e.g., the one that “opened” it) has load at least α . In both cases, the total size of demands assigned by the algorithm to the edges of these cliques is at least

$$\begin{aligned} (c_{\geq L} - 1) \cdot \binom{L}{2} \cdot \alpha &= (c_{\geq L} - 1) \cdot \frac{L}{4} \cdot \min \left\{ 1, \frac{(L-1)B}{L} \right\} \\ &\geq (c_{\geq L} - 1) \cdot \frac{L}{4} \left(1 - \frac{1}{L} \right) = m_{\geq L} \cdot \frac{1}{4} \left(1 - \frac{1}{L} \right) - O(L), \end{aligned} \quad (1)$$

yielding roughly average value $\frac{1}{4}$ from each node of these cliques, as claimed.

Cliques for I_k , for $k \geq 3$. Consider any clique for I_k except the last one to be opened. All edges of this clique have some demand from I_k assigned to it; given the lower limit for this interval, this means that the algorithm has assigned to each such clique total size at least

$$\binom{k}{2} \cdot \min \left\{ \frac{1}{k}, \frac{B}{k+1} \right\} = \frac{k}{2} \cdot \min \left\{ \frac{k-1}{k}, \frac{B(k-1)}{k+1} \right\}.$$

Since $k \geq 3$ and $B \geq 1$, the right-hand side is at least $\frac{k}{4}$. Letting again c_k denote the number of cliques for I_k that the algorithm opens and m_k the corresponding number of nodes/machines, we can count the total value of all but the last I_k clique and we see that the algorithm has assigned to them total size at least

$$(c_k - 1) \cdot \frac{k}{4} = m_k \cdot \frac{1}{4} - O(k). \quad (2)$$

Cliques for I_2 . (Recall that there is no $k = 1$, so this is the last case to consider.) Given the lower limit of the interval I_2 , each I_2 clique (which being a 2-clique is just an edge) has a demand of size at least $\min\{\frac{1}{2}, \frac{B}{3}\}$ assigned to it. So the algorithm assigns total size at least $m_2 \cdot \min\{\frac{1}{4}, \frac{B}{6}\}$ to these I_2 cliques, where m_2 is the number of nodes in these cliques.

Total value of Alg. Since we assumed that the algorithm stops at some point, it means that it could not open more cliques. This means that all but at most $L - 1$ nodes belong to one such clique (the worst case is that it tried to open an L -clique but could not), so $m_{\geq L} + \sum_{k=3}^{L-1} m_k + m_2 \geq m - L$. Then adding the above estimates for the values obtained on each type of clique, we see that the algorithm gets total value at least

$$\begin{aligned} \text{Alg} &\geq \frac{1}{4} \left(1 - \frac{1}{L} \right) \cdot (m - m_2 - L) - O(L^2) + m_2 \cdot \min \left\{ \frac{1}{4}, \frac{B}{6} \right\} \\ &= \frac{1}{4} \cdot (m - m_2) + m_2 \cdot \min \left\{ \frac{1}{4}, \frac{B}{6} \right\} - O(m^{2/3}) \end{aligned}$$

where the last inequality uses the fact that $L = m^{1/3}$.

Notice that if the minimum in the last line is $\frac{1}{4}$, then we obtain $\text{Alg} \geq (\frac{1}{2} - O(\frac{1}{m^{1/3}}))\text{OPT}$ as desired (recall $\text{OPT} \leq \frac{m}{2}$ since each machine has Nominal capacity 1 and each demand is assigned to two machines). **So assume this is not the case, namely $B < \frac{3}{2}$.** Under this assumption

$$\text{Alg} \stackrel{\text{with ass.}}{\geq} \frac{1}{4} \cdot (m - m_2) + \frac{B}{6} \cdot m_2 - O(m^{2/3}) \quad (3)$$

Value of OPT. We analyze OPT again under the assumption $B < \frac{3}{2}$. The Failover constraints also ensure that in order to accommodate the demand from I_2 in case of failure, any node that receives a demand from I_2 can have total size assigned to it at most

$$B - \min \left\{ \frac{1}{2}, \frac{B}{3} \right\} \stackrel{\text{with ass.}}{=} \frac{2B}{3},$$

the last equation due to the assumption $B < \frac{3}{2}$. For all other nodes, OPT can assign at most size 1 per node due to the Nominal capacity constraint. Let m_2^{OPT} be the number of nodes where OPT schedules a demand from I_2 . Again, since the size of each demand is counted towards the Nominal capacity of two nodes, the total size scheduled by OPT is

$$\text{OPT} \leq \frac{1}{2} \left(m_2^{\text{OPT}} \cdot \frac{2B}{3} + (m - m_2^{\text{OPT}}) \cdot 1 \right) = \frac{1}{2} \cdot (m - m_2^{\text{OPT}}) + \frac{B}{3} \cdot m_2^{\text{OPT}} \quad (4)$$

Notice that since every demand in I_2 has size $> \min\{\frac{1}{2}, \frac{B}{3}\} \geq \frac{1}{3}$, the Failover constraints ensure that in OPT (as well as in our algorithm) the demands from I_2 that are scheduled form a matching, i.e. no 2 such demands can share a node/machine. So m_2^{OPT} (resp. m_2) is just twice the number of I_2 demands scheduled by OPT (resp. our algorithm). Moreover, both Alg and OPT schedule a prefix of the instance. Since OPT gets at least as much value as Alg, it means that it scheduled a prefix that is at least as long; in particular it schedules at least as many I_2 demands as our algorithm. Together these observations imply that $m_2^{\text{OPT}} \geq m_2$. Then given inequalities (3) and (4), under the assumption $B < \frac{3}{2}$ we obtain that $\text{Alg} \geq (\frac{1}{2} - O(\frac{1}{m^{1/3}}))\text{OPT}$ as desired. This concludes the proof of Theorem 1.

3 Sublinear Additive Regret in the Stochastic Model

We now consider FAILOVER in the online stochastic model, where, instead of being adversarial, the size S_t of each demand now comes independently from an unknown distribution μ over $[0, \min\{1, \frac{B}{2}\}]$. Again, at time t the algorithm observes the size S_t of the current demand and irrevocably assigns it to two of the m machines. We still use $\text{OPT} = \text{OPT}(S_1, \dots, S_n)$ to denote the value of (sum of the sizes scheduled by) the optimal strategy, which is now a random quantity.

Our main result is algorithm FailoverStochastic that achieves a sublinear additive loss compared to OPT in this stochastic model (restated from the introduction for convenience).

Theorem 2. *For the FAILOVER problem in the stochastic arrival model, there is a poly-time algorithm that achieves utilization $\text{OPT} - O(\text{OPT}^{5/6} \log \text{OPT})$ with probability $1 - O(\frac{1}{m})$.*

The algorithm relies on a learn-and-pack approach that uses previously seen items to compute a *template* for packing the next items. This process is performed in rounds. Each round starts by assigning the first demand of the round on a pair of (empty) machines. Then, we

iteratively create a template for the first $n_k := 2^k$ items of the round, which we use to schedule the next n_k items. When the number of machines needed for the template (along with some slack) exceeds the number of available machines, the current round terminates and the next round begins. The next round maintains no knowledge of the previous demands; it only takes as input the number of empty machines \tilde{m} which it is allowed to use. A schematic overview of this process is presented in Figure 4.

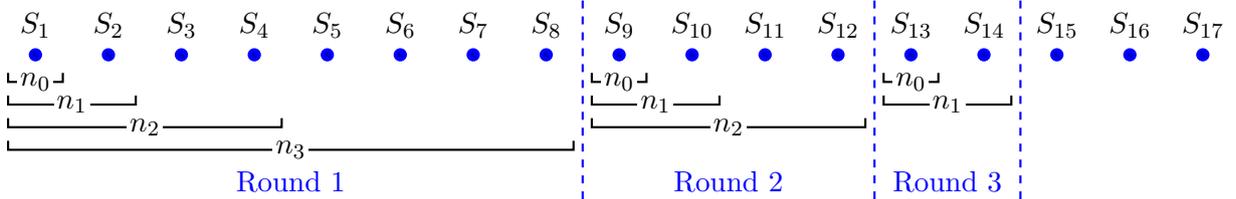


Figure 4: Schematic overview of algorithm FailoverStochastic.

Before describing the algorithm in more detail, an important question that arises is how to use the templates to schedule the future demands. A crucial component in this process are *monotone matchings*, which only match two values if the second is at least as big as the first.

Definition 1 (Monotone matching). *Given two sequences $x_1, \dots, x_n \in \mathbb{R}$ and $y_1, \dots, y_n \in \mathbb{R}$, a monotone matching π from the x_t 's to the y_t 's is an injective function from a subset $I \in \{1, \dots, n\}$ to $\{1, \dots, n\}$ such that $x_i \leq y_{\pi(i)}$ for all $i \in I$. We say that x_i is matched to $y_{\pi(i)}$ if $i \in I$, and x_i is unmatched otherwise.*

Monotone matchings will allow us to match future demands (x_i 's) to the demands that are part of a template ($y_{\pi(i)}$'s) and put the former in the place of the latter (since $x_t \leq y_{\pi(i)}$). A surprising result of Rhee and Talagrand [19] is that if the two sequences are sampled i.i.d. from the same distribution, then almost all items can be matched, and moreover such a matching can be found online (see the paper for a more general result where the sequences may come from different distributions).

Theorem 3 (Monotone Matching Theorem [19]). *Suppose the random variables A_1, \dots, A_n and B_1, \dots, B_n are all sampled independently from a distribution μ . Then there is a constant cst such that with probability at least $1 - e^{-cst \cdot \log^{3/2} n}$ there is a monotone matching π of the A_i 's to the B_i 's where at most $cst \cdot \sqrt{n} \log^{3/4} n$ of the A_i 's are unmatched. Moreover, this matching can be computed even if the sequence A_1, \dots, A_n is revealed online.*

3.1 Algorithm

We are now ready to present the details of the FailoverStochastic algorithm.

FailoverStochastic: The algorithm just repeatedly calls the procedure OneRound below, passing to it the number of machines that are still available/unopened (e.g. initially it calls OneRound(m)); it does this for $\frac{\log m}{\log 4/3}$ rounds.

OneRound(\tilde{m}): This procedure receives as input the number \tilde{m} of machines that it is allowed to open. It is convenient to rename the demands and use Y_t to denote the t th demand seen by OneRound (which are still sampled i.i.d. from μ). Similar to the work of Rhee and Talagrand [19], this algorithm works in phases: As mentioned earlier, each phase k sees the previous $n_k = 2^k$ items and creates a template based on them, which will then be used to schedule the next n_k

items. To create this template, we define the offline problem `OFFMINFAILOVER` of minimizing the number of machines that are required to schedule these n_k items. To solve this problem, we design an approximation algorithm `OffMinFailoverAlg` in [Section 4](#) achieving a sublinear approximation guarantee. Specifically, let $\overline{\text{OPT}}_{mach}(x_1, \dots, x_n)$ be the number of machines that `OffMinFailoverAlg` (with $\varepsilon = 1/n_k^{1/6}$) uses to schedule the demands x_1, \dots, x_n . `OneRound` is then as follows:

OneRound: Given a number of available machines \tilde{m} :

- 1: Assign the first demand Y_1 to an empty edge by itself, opening 2 machines.
- 2: For phases $k = 0, 1, 2, \dots$
 - (a) See the first n_k items Y_1, \dots, Y_{n_k} . Run the algorithm `OffMinFailoverAlg` from [Section 4](#) (with $\varepsilon = 1/n_k^{1/6}$) to find a solution for them that uses $\overline{\text{OPT}}_{mach}(Y_1, \dots, Y_{n_k})$ machines; let $templ(t)$ denote the pair of machines that Y_t is assigned to. This solution is our template.

(b) If

$$\#\{\text{already open machines}\} + \underbrace{\overline{\text{OPT}}_{mach}(Y_1, \dots, Y_{n_k})}_{\text{machines from template}} + \underbrace{cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k}_{\text{predicted unmatched demand}} + 2m^{5/6} > \tilde{m},$$

then STOP.

- (c) Else, open a clique of $\overline{\text{OPT}}_{mach}(Y_1, \dots, Y_{n_k})$ machines. Upon the arrival of each of the next n_k demands $Y_{n_k+1}, \dots, Y_{2n_k}$, assign them to machines based on the template. More precisely, find the Rhee-Talagrand monotone matching π guaranteed by [Theorem 3](#) from the new to the old demands (as the new ones arrive online). Schedule each matched new demand Y_t to the pair of machines that $Y_{\pi(t)}$ occupied in the template, namely the machine pair $templ(\pi(t))$. For each unmatched new demand, schedule it on an edge by itself (opening two more machines for each). If at any point the execution tries to open more than \tilde{m} machines, declare FAIL.
-

3.2 Analysis

We next discuss the main ideas for the analysis of the algorithm `FailoverStochastic`, leading to the proof of [Theorem 2](#). We assume throughout that m is at least a sufficiently large constant, else the success probability $1 - O(\frac{1}{m})$ trivially holds.

We need to develop two important components for the analysis that are done in their own sections. To at least state them, let $\text{OPT}_{mach}(J)$ denote the minimum number of devices needed to assign all demands from set J satisfying the Nominal and Failover constraints.

First component ([Section 4](#)): The first component is the aforementioned algorithm `OffMinFailoverAlg` that is called within `OneRound`. It relies on a novel configuration LP, (LP_{mach}), and a post-processing algorithm to realize a rounded LP solution as a feasible assignment. It has the following guarantee:

Theorem 4. *There exists a poly-time algorithm, `OffMinFailoverAlg`, that given $\varepsilon \in (0, 1)$, finds a solution for `OFFMINFAILOVER` with at most $(1 + O(\varepsilon))\text{LP}_{mach} + O(\frac{1}{\varepsilon^5}) \leq (1 + O(\varepsilon))\text{OPT}_{mach} + O(\frac{1}{\varepsilon^5})$ machines.*

Choosing ε appropriately, we will be able to create a template using at most $\mathbb{E} \text{OPT}_{mach}(Y_1, \dots, Y_{n_k}) + o(n_k)$ devices in expectation for the next n_k arrivals.

Second component ([Section 5](#)): Recall from the technical overview that a worrisome aspect of `FailoverStochastic` is that each call to `OneRound` does not re-use machines from previous

rounds. To show this is not too wasteful, we show that $\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T)$ is approximately linear in T . We do so by giving a quantitative convergence theorem of $\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T)$ to $T \cdot c(\mu)$, where $c(\mu)$ is a constant that characterizes the “average number of devices needed per demand.” That is, we show the following:

Theorem 5. *Let μ be a distribution supported on $[0, \min\{1, \frac{B}{2}\}]$. Then there exists a scalar $c(\mu)$ such that for every $T \in \mathbb{N}$, we have*

$$\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) \in T \cdot c(\mu) \pm O(T^{5/6}),$$

where X_1, \dots, X_T are i.i.d. samples from μ .

Thus splitting the first $2n_k$ demands into two rounds of n_k demands each costs us only an extra $o(n_k)$ devices.

With those two results in hand, the core of the analysis is that **OneRound** gets good value density, i.e., the ratio of value over number of machines m . We use $\mathbb{E}S_0$ to denote the expected value of the size of a demand (which is the same as $\mathbb{E}S_t$ for any t).

Specifically, according to **Theorem 5**, there is a scalar $c(\mu)$ such that **OPT** is able to fit roughly $\frac{1}{c(\mu)}$ demands per machine. Each such demand gives value roughly $\mathbb{E}S_0$; so the intuition is that the best possible density value/machine should be around $\frac{\mathbb{E}S_0}{c(\mu)}$. We first make this formal in the next lemma.

Lemma 1. *With probability at least $1 - \frac{2}{m^2}$ we have*

$$\text{OPT} \leq m \cdot \frac{\mathbb{E}S_0}{c(\mu)} + O(m^{5/6}).$$

Crucially, the next lemma says that **OneRound** almost achieves this density.

Lemma 2. *Let $Open$ be the number of machines opened by **OneRound**(\tilde{m}) (which is a random variable). Then with probability at least $1 - \frac{1}{m^2}$, the total value of the demands scheduled by **OneRound**(\tilde{m}) is at least*

$$\text{value of OneRound}(\tilde{m}) \geq \frac{\mathbb{E}S_0}{c(\mu)} \cdot Open - O(m^{5/6}).$$

Given this lemma, we see that the total value of the **FailoverStochastic** algorithm (which repeatedly calls **OneRound**) is approximately $\frac{\mathbb{E}S_0}{c(\mu)}$ times the total machines opened during the execution. By showing that the number of machines **FailoverStochastic** opens is $\approx m$, we then almost match the upper bound on **OPT** from **Lemma 1**.

Lemma 3. *With probability $1 - O(\frac{1}{m})$, **FailoverStochastic** opens at least $m - 5cst_5 \cdot m^{5/6}$ machines (where cst_5 is the constant from **Lemma 14**).*

These lemmas quickly lead to the proof of **Theorem 2**.

*Proof of **Theorem 2**.* Let $L := \frac{\log m}{\log 4/3}$ denote the number of calls to **OneRound** that **FailoverStochastic** makes, and let val_i and $Open_i$ be the value obtained and number of machines opened by the i -th call. Employing **Lemma 2** on these L calls, we have that with probability at least $1 - \frac{L}{m^2}$ the total value of **FailoverStochastic** is

$$\text{algo value} = val_1 + \dots + val_L \geq \frac{\mathbb{E}S_0}{c(\mu)} \cdot \sum_{i \leq L} Open_i - O(m^{5/6} \log m).$$

Moreover, from [Lemma 3](#), with probability at least $1 - O(\frac{1}{m})$ the total number of machines open $\sum_{i \leq L} \text{Open}_i$ is at least $m - 5cst_5 \cdot m^{5/6}$, in which case we get

$$\text{algo value} \geq m \cdot \frac{\mathbb{E}S_0}{c(\mu)} - O(m^{5/6} \log m). \quad (5)$$

Furthermore, from [Lemma 1](#) we have that $\text{OPT} \leq m \cdot \frac{\mathbb{E}S_0}{c(\mu)} + O(m^{5/6})$ with probability at least $1 - \frac{2}{m^2}$. So by taking a union bound and combining this with the above lower bound on the algorithm's value, we get that with probability $1 - O(\frac{1}{m})$

$$\text{algo value} \geq \text{OPT} - O(m^{5/6} \log m).$$

Since (5) also implies that $\text{OPT} \geq \Omega(m)$, the previous bound is at least $\text{OPT} - O(\text{OPT}^{5/6} \log \text{OPT})$. This concludes the proof of [Theorem 2](#). \square

We conclude this section by proving the lower bound on the value density of [OneRound](#) from [Lemma 2](#). We defer the proofs of [Lemma 1](#) and [3](#) to [Appendix D](#).

3.2.1 Proof of [Lemma 2](#)

First, we control in high-probability the number of phases that [OneRound](#)(\tilde{m}) executes before stopping or failing; this will be important to avoid dependencies on the total number of demands n in the instance, which can be arbitrarily bigger than the scale of the effective instance.

Claim 1. *With probability at least $1 - \frac{1}{m^3}$ the algorithm [OneRound](#) performs at most*

$$\bar{k} := \log \left(\frac{\tilde{m}}{c(\mu)} + O(\tilde{m}^{5/6}) + 3 \log^{\frac{3}{2}} m \right) \quad (6)$$

phases.

Proof. Recall that the demands sizes Y_1, Y_2, \dots that [OneRound](#) sees are still i.i.d. samples from the original distribution μ . In [Lemma 13](#) (with $m = \tilde{m}$ and $\delta = \frac{1}{m^3}$) in the appendix we show that with probability at least $1 - \frac{1}{m^3}$ [OneRound](#) can schedule at most $\frac{\tilde{m}}{c(\mu)} + O(\tilde{m}^{5/6}) + 3 \log^{\frac{3}{2}} m$ many of these demands; for some intuition, [Lemma 1](#) indicates that even OPT cannot schedule more than roughly these many demands. Since this quantity is exactly $n_{\bar{k}}$, [OneRound](#) cannot complete phase \bar{k} (there are $2n_{\bar{k}}$ demands by the end of it) and the claim holds. \square

Next, we need to bound how many machines are opened by [OneRound](#), which in particular affects the probability of it failing. For a phase k , let $M_k := \overline{\text{OPT}}_{\text{mach}}(Y_1, \dots, Y_{n_k})$ denote the number of machines in the template solution, and let U_k be the number of additional machines that had to be open to accommodate the unmatched demands among $Y_{n_k+1}, \dots, Y_{2n_k}$, namely twice the number of unmatched items. Notice that these quantities are well defined even for phases that the algorithm did not execute. The quantity $M_k + U_k$ is then the number machines that the algorithm [OneRound](#) opens in phase k (if it executes it). We have the following bounds for the number of machines open, at least for a phase k where the number of items n_k is sufficiently large (but still sublinear in m).

Claim 2. *Let $k_0 := (\frac{2}{cst_2} \log m)^{2/3}$ for a sufficiently small constant cst_2 . Then there is a constant cst_1 such that:*

1. *For $k \geq k_0$, we have $M_k \in n_k \cdot c(\mu) \pm cst_1 \cdot n_k^{5/6}$ with probability $\geq 1 - \frac{1}{m^3}$*
2. *For $k \geq k_0$, we have $U_k \leq cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k$ with probability $\geq 1 - \frac{1}{m^3}$*

3. $n_{k_0} \leq m^{5/6}$.

Proof. Consider a phase $k \geq k_0$. Since the demand sizes Y_1, \dots, Y_{2n_k} seen in this phase are i.i.d. samples from the original distribution μ , we can bound the minimum number of machines $\text{OPT}_{mach}(Y_1, \dots, Y_{n_k})$ (using [Corollary 1](#) in [Appendix D.1](#) with $\lambda = n_k^{1/3}$)

$$\text{OPT}_{mach}(Y_1, \dots, Y_{n_k}) \in n_k \cdot c(\mu) \pm O(n_k^{5/6})$$

with probability at least $1 - 2e^{-\frac{n_k^{2/3}}{2}}$. Moreover, employing the guarantee of the algorithm `OffMinFailoverAlg` used to build the template ([Theorem 4](#) with $\varepsilon = 1/n_k^{1/6}$), we get

$$M_k \in n_k \cdot c(\mu) \pm cst_1 \cdot n_k^{5/6}$$

with probability at least $1 - 2e^{-\frac{n_k^{2/3}}{2}}$ for some constant cst_1 . But since $n_k = 2^k \geq 2^{k_0}$, a quick calculation shows that this probability is at least $1 - \frac{1}{m^3}$, proving the first item of the claim.

To control U_k , we can use the Monotone Matching Theorem ([Theorem 3](#)) with the first sequence of sizes being the demands from the template, i.e., $(B_1, \dots, B_{n_k}) = (Y_1, \dots, Y_{n_k})$, and the second one being the demands that we attempted to match to them, namely $(A_1, \dots, A_{n_k}) = (Y_{n_k+1}, \dots, Y_{2n_k})$ to obtain that the number of unmatched demands it at most $cst \cdot \sqrt{n_k} \log^{3/4} n_k$ with probability at least $1 - e^{-cst \cdot \log^{3/2} n_k}$, and hence with this probability

$$U_k \leq 2cst \cdot \sqrt{n_k} \log^{3/4} n_k.$$

Again because $k \geq k_0$, we get that this probability is at least $1 - \frac{1}{m^3}$, proving Item 2 of the claim (by taking $cst_1 \geq 2cst$ we can just replace the latter by the former).

The last item $n_{k_0} \leq m^{5/6}$ of the claim can be directly verified using the fact that we assumed m is at least a sufficiently large constant. \square

Recall that `OneRound` only fails when the number of machines $M_k + U_k$ actually opened in a phase is bigger than it “predicted” in [Line 2.\(b\)](#), and this prediction is exactly M_k plus the upper bound U_k from [Claim 2](#) plus a slack of $2m^{5/6}$. By considering all phases, it is now easy to upper bound the probability that `OneRound` fails (\bar{k} is defined in [\(6\)](#)).

Claim 3. *The probability that `OneRound` fails is at most $\frac{\bar{k}+1}{m^3}$.*

Proof. Fix any phase k , and we claim that the probability that `OneRound` fails on this phase is at most $\frac{1}{m^3}$. If `OneRound` fails on phase k , then it did not STOP in [Line 2.\(b\)](#), so

$$\# [\text{machines open before phase } k] + \overline{\text{OPT}}_{mach}(Y_1, \dots, Y_{n_k}) + cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k + 2m^{5/6} \leq \tilde{m},$$

but it ran out of machines during phase k , namely

$$\# [\text{machines open before phase } k] + (M_k + U_k) > \tilde{m}.$$

Since $M_k = \overline{\text{OPT}}_{mach}(Y_1, \dots, Y_{n_k})$, these observations imply that $U_k > cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k + 2m^{5/6}$. This is impossible if $n_k \leq m^{5/6}$, because the number of machines U_k opened for the unmatched demands is at most twice the number n_k of demands considered for the matching. So we must have $n_k > m^{5/6}$ (and so from [Claim 2](#) $k \geq k_0$) and at least $U_k > cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k$; but again by [Claim 2](#) the latter happens with probability at most $\frac{1}{m^3}$. Thus, the probability that `OneRound` fails on phase k is at most $\frac{1}{m^3}$.

Moreover, by [Claim 1](#), with probability at least $1 - \frac{1}{m^3}$ `OneRound` has at most \bar{k} phases. Then taking a union bound, we see that the event that `OneRound` has at most \bar{k} phases and in all of them it does not fail holds with probability at least $1 - \frac{\bar{k}+1}{m^3}$; in particular, with at least this much probability the algorithm does not fail in its execution, which proves the claim. \square

We now finally lower bound the value that **OneRound** gets. Let τ be the (random) index of the last phase attempted by **OneRound**, namely where Line 2.(c) is executed. As long as it does not fail on the last phase τ (which by the previous claim happens with probability at least $1 - \frac{\bar{k}+1}{m^3}$) **OneRound** gets the value of all items up until this phase, that is

$$\text{value of OneRound} \geq Y_1 + \dots + Y_{2n_\tau} \geq Y_1 + \dots + Y_{2n_{\min\{\tau, \bar{k}\}}}. \quad (7)$$

Recall that the Y_i 's are independent and each has mean $\mathbb{E}S_0$. Then employing the Chernoff bound (Lemma 11) with $\lambda = \sqrt{n_{\bar{k}} \log(m^3 \cdot n_{\bar{k}})}$, for any fixed $t \leq n_{\bar{k}}$ we have that

$$Y_1 + \dots + Y_t \geq t \cdot \mathbb{E}S_0 - \sqrt{n_{\bar{k}} \log(m^3 \cdot n_{\bar{k}})} \quad \text{with probability at least } 1 - \frac{1}{m^3 \cdot n_{\bar{k}}}.$$

Then taking a union bound over (7), the previous displayed inequality for all $t \leq n_{\bar{k}}$, and over the event that **OneRound** has at most \bar{k} phases (which holds with probability at least $1 - \frac{1}{m^3}$) we get that

$$\begin{aligned} \text{value of OneRound} &\geq 2n_{\min\{\tau, \bar{k}\}} \cdot \mathbb{E}S_0 - \sqrt{n_{\bar{k}} \log(m^3 \cdot n_{\bar{k}})} \\ &= 2n_\tau \cdot \mathbb{E}S_0 - \sqrt{n_{\bar{k}} \log(m^3 \cdot n_{\bar{k}})} \\ &\geq 2n_\tau \cdot \mathbb{E}S_0 - O(m^{5/6}) \quad \text{with probability } \geq 1 - \frac{\bar{k} + 3}{m^3}. \end{aligned} \quad (8)$$

To conclude the proof of Lemma 2 we just need to relate this quantity to the number of machines opened by **OneRound**. Let $Open_\ell$ be the number of machines opened until (including) phase ℓ , and recall that $Open$ is the number of machines opened over all phases. Since the number of machines opened on phase k is $M_k + U_k$ (plus two machines for the first demand Y_1), we have

$$Open_\ell = 2 + (M_1 + U_1) + \dots + (M_\ell + U_\ell) \quad (9)$$

To upper bound the right-hand side, for the phases $k < k_0$ we just use the fact that $M_k + U_k \leq 2n_k + 2n_k = 4n_k$, since both in the template and for the unmatched demands we never open more than 2 machines per demand considered (and n_k demands are considered in each part). For each phase $k = k_0, \dots, \bar{k}$ we can use Claim 2 to upper bound $M_k + U_k$ with probability at least $1 - \frac{2}{m^3}$ by

$$\begin{aligned} M_k + U_k &\leq n_k \cdot c(\mu) + cst_1 \cdot n_k^{5/6} + cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k \\ &\leq n_k \cdot c(\mu) + cst_3 \cdot n_k^{5/6} \end{aligned}$$

for some constant cst_3 . Together these bounds give that with probability at least $1 - \frac{2\ell}{m^3}$

$$Open_\ell \leq 2 + \sum_{k < k_0} 4n_k + \sum_{k=k_0}^{\ell} \left(n_k \cdot c(\mu) + cst_3 \cdot n_k^{5/6} \right).$$

To further upper bound the first summation on the right-hand side, because of the exponential relationship $n_k = 2^k$, we have $\sum_{k < k_0} 4n_k \leq 8n_{k_0-1} \leq O(m^{5/6})$, the last inequality coming from Claim 2; for the second summation, we analogously have $\sum_{k=k_0}^{\ell} n_k \leq 2n_\ell$ and $\sum_{k=k_0}^{\ell} n_k^{5/6} \leq O(n_\ell^{5/6})$. Therefore,

$$Open_\ell \leq 2n_\ell \cdot c(\mu) + O(n_\ell^{5/6}) + O(m^{5/6}) \quad \text{with probability at least } 1 - \frac{2\ell}{m^3}. \quad (10)$$

Finally, since by [Claim 1](#) the number of phases τ performed by `OneRound` is at most \bar{k} with probability at least $1 - \frac{1}{m^3}$, the total number of machines open can be upper bounded

$$Open \leq Open_{\min\{\tau, \bar{k}\}} \leq 2n_\tau \cdot c(\mu) + O(n_{\bar{k}}^{5/6}) + O(m^{5/6}) \leq 2n_\tau \cdot c(\mu) + O(m^{5/6})$$

with probability at least $1 - \frac{2\bar{k}+1}{m^3}$.

Finally, taking a union bound to combine this inequality with [\(8\)](#), we get that

$$\text{value of OneRound} \geq \frac{\mathbb{E}S_0}{c(\mu)} \cdot Open - O(m^{5/6})$$

with probability at least $1 - \frac{3\bar{k}+4}{m^3}$. Since m is at least a sufficiently large constant, we have $m \geq 3\bar{k} + 4$, and the bound from the displayed inequality holds with probability at least $1 - \frac{1}{m^2}$. This finally concludes the proof of [Lemma 2](#).

4 Offline Machine Minimization

In this section we consider the aforementioned (offline) minimization version of `FAILOVER`, which we call `OFFMINFAILOVER`: Given a failover capacity $B \geq 1$ and a collection of demands such that demand j has size $s_j \in [0, \min\{1, \frac{B}{2}\}]$, we need to assign *all demands* to pairs of machines while satisfying the Nominal and Failover constraints, and the goal is to minimize the number of machines used. As before, we use $\text{OPT}_{mach} = \text{OPT}_{mach}(s_1, \dots, s_n)$ to denote the cost of (i.e. number of machines in) the optimal solution.

The main result of this section ([Theorem 4](#), restated) is an efficient algorithm with a sublinear additive regret for this problem (when ε is set appropriately). We remark that a sublinear regret (compared to, say, a constant approximation) is necessary due to its use in [Section 3](#). In fact, the algorithm compares against the stronger optimum of an LP relaxation for the problem (denoted by (LP_{mach}) , and defined below), which will be crucially used in [Section 5](#). We let LP_{mach} denote the optimal value of this LP.

Theorem 4. *There exists a poly-time algorithm, `OffMinFailoverAlg`, that given $\varepsilon \in (0, 1)$, finds a solution for `OFFMINFAILOVER` with at most $(1 + O(\varepsilon))\text{LP}_{mach} + O(\frac{1}{\varepsilon^5}) \leq (1 + O(\varepsilon))\text{OPT}_{mach} + O(\frac{1}{\varepsilon^5})$ machines.*

As hinted above, our algorithm is based on converting a solution of a configuration LP into a good assignment of demands to pairs of machines. But crucially, while the configuration of each machine controls the total size of demands serviced by it, it has no information how these demands are distributed over the “edges” incident to the machine, which is important for adequately handling the Failover constraints. The post-processing of the LP solution is the one in charge of creating a feasible (and low-cost) assignment from this limited control offered by the LP.

4.1 Configuration LP

Consider an assignment of the demands into some number of machines. We can view the collection of demands assigned to (the edges incident to) a given machine as a configuration. Precisely, we define a *configuration* C to be a subset of the demands such that $\sum_{s \in C} s \leq 1$ and $\sum_{s \in C} s + \max_{s \in C} s \leq B$. Note that the first constraint is exactly the Nominal constraint, while the second is a relaxation of the Failover constraint, because the most-loaded edge incident on some machine can be larger than the single largest demand assigned to that machine. Thus,

our notion of configuration does not take in to account how the demands are assigned to the respective edges incident on each machine.

To define our configuration LP, we suppose the input collection of demands is partitioned into T demand types such that type t consists of n_t -many demands each with size s_t . Thus each configuration C can be represented by a number $n_t(C) \in \mathbb{N}$ of demands for each type t such that $\sum_t n_t(C) \cdot s_t \leq 1$ and $\sum_t n_t(C) \cdot s_t + \max_{t | n_t(C) > 0} s_t \leq B$. We are ready to define our configuration LP:

$$\begin{aligned} \min \quad & \sum_C x_C \\ \text{s.t.} \quad & \sum_C n_t(C) \cdot x_C \geq 2n_t \quad \forall t \\ & x \geq 0 \end{aligned} \tag{LP}_{mach}$$

Note that the definition of (LP_{mach}) depends on how the demands are partitioned into types. We show in [Appendix E.1](#) that the optimal value of (LP_{mach}) does not depend on the particular type partition. Thus, throughout the analysis, we will use whichever type partition is convenient (unless a particular one is specified).

It is immediate that (LP_{mach}) is a relaxation of OFFMINFAILOVER by taking the natural setting of the x -variables defined by a feasible assignment to machines: just let x_C be the number of machines whose collection of demand sizes assigned to its edges are exactly those in C . In particular, we have that $\text{LP}_{mach} \leq \text{OPT}_{mach}$.

Although (LP_{mach}) has exponentially many variables in general, we can approximately solve it via column generation similar to the standard bin packing configuration LP [[10](#), [20](#)] (see [Appendix E.2](#)).

Lemma 4. *We can find in poly-time an extreme point solution of (LP_{mach}) with objective value at most $\text{LP}_{mach} + 1$.*

Further, observe that (LP_{mach}) only has T non-trivial constraints, so by the standard rank argument (see for example Lemma 2.1.3 of [[12](#)]) any extreme point solution of (LP_{mach}) has at most T non-zero variables. Thus, the next lemma follows immediately by rounding up all the fractional variable of an extreme point solution.

Lemma 5. *Given an extreme point of (LP_{mach}) with objective value Val , rounding up all fractional variables to the next largest integer gives an integral solution to (LP_{mach}) with objective value at most $Val + T$.*

To summarize this section, we can efficiently obtain a collection of configurations, each corresponding to a machine, that “covers” all the demands. However, these configurations do not specify how to actually assign the demands to the edges incident on the corresponding machine. This is the goal of the next section.

4.2 Matching configurations

We say that a collection \mathcal{C} of configurations is feasible if it comes from an integer solution for (LP_{mach}) , i.e. setting x_C to be the number of times C appears in \mathcal{C} gives a feasible solution for (LP_{mach}) . Our goal in this section is to realize such collection by actually assigning demands to edges. The main challenge is satisfying the actual Failover constraints.

For simplicity assume $\sum_{C \in \mathcal{C}} n_t(C) = 2n_t$ for all types t , i.e. each demand appears on exactly 2 configurations (drop from the configurations what is extra). We can think of \mathcal{C} (with, say, N configurations) as a graph on N nodes/machines, where node/machine $C \in \mathcal{C}$ has $n_t(C)$ “slots” for demands of type t . While this gives the right number of slots $2n_t$ to accommodate the demands of each type t , we still need to specify to which edge (pair of machines) each of the n_t

demands of type t is assigned in a way that satisfies the Nominal and Failover constraints. (We can alternatively see this as a graph realization problem: each node C as having a requirement $n_t(C)$ of “edges of type t ” (which we call its t -degree) and we want to create edges of different types (i.e., assignment of demands to pairs of nodes) to satisfy these requirements while also satisfying the Nominal and Failover constraints.)

To see the challenge, consider a fixed node/configuration C . Regardless of how we assign demands to edges (as long as it is consistent with the slots of the configurations), the Nominal constraint of C is satisfied: it will receive total size $\sum_t n_t(C) \cdot s_t = \sum_{s \in C} s$, which is at most 1 by definition of a configuration. This is not the case for the Failover constraint. This is again because the definition of configuration only gives us the relaxed version of the Failover constraint $\sum_{s \in C} s + \max_{s \in C} s \leq B$. In particular, the blue term only considers the largest demand assigned to machine C instead of the most-loaded edge incident to C . However, these two quantities are the same *if we are able to assign at most one demand per edge*. (In the graph realization perspective, it means that it suffices to construct a *simple* graph with the desired t -degrees.) But it is not clear that such an assignment should even exist, let alone be found efficiently.

The main result of this section is that – by opening slightly more machines – we can find such an assignment that realizes any given collection of configurations satisfying both the Nominal and Failover constraints.

Theorem 6. *Consider an instance of OFFMINFAILOVER with T demand types. Given a collection \mathcal{C} of N configurations that is feasible for $(\text{LP}_{\text{mach}})$, we can find in poly-time a feasible solution for OFFMINFAILOVER that uses at most $N + O(DT)$ machines, where D is the maximum number of demands in any configuration in \mathcal{C} .*

For that, we will need the following subroutine to assign some demands outside of their respective configurations.

Lemma 6. *There is a poly-time algorithm for OFFMINFAILOVER that uses at most $8 \cdot S + 2$ machines, where S is the sum of the size of the demands in the instance.*

Proof. Our algorithm will only open edges – that is, we will open machines in pairs and will only assign demands to the edges of the paired machines. Our algorithm is the following: Consider the demands in any order. Assign each demand to already-opened edge as long as the Nominal and Failover constraints remain satisfied. Else open a new edge and assign the demand there.

It is clear that the algorithm is efficient and satisfies the constraints. We claim this algorithm opens at most $8 \cdot S + 2$ machines (i.e. $4 \cdot S + 1$ edges). To see this, note that every edge except at most one has load at least $\frac{1}{4}$. If not, then consider the first time that there are two open edges with load less than $\frac{1}{4}$. It must be the case that the last demand s_j the algorithm considered up to this point had size less than $\frac{1}{4}$, but the algorithm decided to open a new edge for this demand rather than assign it to a previous edge (u, v) that already had load less than $\frac{1}{4}$. However, assigning demand s_j to the edge (u, v) is feasible; the left-hand side of the Failover constraint for machine u , say, would be at most $\frac{1}{2}$ (total size assigned to machine u) plus $\frac{1}{2}$ (total size assigned to the edge (u, v) , the only one incident to u), which is at most the Failover capacity $B \geq 1$.) This contradicts the definition of the algorithm. \square

The algorithm guaranteed by [Theorem 6](#) is the following. In order to simplify the notation, as before we assume without loss of generality that \mathcal{C} has $\sum_{C \in \mathcal{C}} n_t(C) = 2n_t$ for all types t .

Proof of Theorem 6. It is clear that `MatchConfigs` runs in polynomial time, and assigns all demands to edges. Further, this assignment satisfies both the Nominal and Failover constraints,

MatchConfigs: Given a collection \mathcal{C} of N configurations:

- 1: Open N machines – one corresponding to each configuration in \mathcal{C} .
 - 2: Consider demand types in arbitrary order $t = 1, \dots, T$.
 - 3: When considering demand type t , partition the collection \mathcal{C} into two collections \mathcal{L}_t and \mathcal{R}_t such that their total t -degrees $\sum_{C \in \mathcal{L}_t} n_t(C)$ and $\sum_{C \in \mathcal{R}_t} n_t(C)$ differ by at most D_t , where D_t is the maximum number of type t demands in any configuration. (This can be achieved, e.g., by initializing $\mathcal{L}_t, \mathcal{R}_t = \emptyset$, and adding configurations one-by-one to the set with minimum total t -degree.)
 - 4: Given this partition, as long as there exists a configuration $C \in \mathcal{L}_t$ that is currently assigned less than $n_t(C)$ demands of type t , we pick such a configuration and assign a demand of type t to an arbitrary edge (C, C') (for $C' \in \mathcal{R}_t$) that has not yet been assigned a demand of any type and such that C' is currently assigned less than $n_t(C')$ demands of type t . If no such edge exists, then we stop and move on to the next demand type.
 - 5: Once we are done considering all demand types, assign all the currently unassigned demands to new machines using [Lemma 6](#).
-

because we assign at most one demand per edge in Step 4 (see discussion in the beginning of this section), and Step 5 guarantees a feasible assignment for the remaining demands.

It remains to show that it opens $N + O(DT)$ machines. In particular, by [Lemma 6](#) it suffices to show that the total size of all unassigned demands that reach Step 5 is $O(DT)$. When considering demand type t , there are two possibilities:

Case 1: Step 4 assigns $n_t(C)$ type t demands to each $C \in \mathcal{L}_t$. In this case it assigns $\sum_{C \in \mathcal{L}_t} n_t(C)$ type t demands to edges between \mathcal{L}_t and \mathcal{R}_t , while the total number of type t demands is

$$n_t = \frac{1}{2} \left(\sum_{C \in \mathcal{L}_t} n_t(C) + \sum_{C \in \mathcal{R}_t} n_t(C) \right) \leq \sum_{C \in \mathcal{L}_t} n_t(C) + \frac{D_t}{2}, \quad (11)$$

where the inequality uses the fact that the t -degree of \mathcal{R}_t is at most that of \mathcal{L}_t plus D_t . Thus, at most $\frac{D_t}{2}$ demands of type t remain unassigned and reach Step 5. The total size of these demands is at most $\frac{1}{2}$, since D_t demands of type t are in a valid configuration. Hence the total size of the unassigned demands of all types is at most $\frac{T}{2} \leq O(DT)$.

Case 2: Step 4 fails to assign $n_t(\bar{C})$ to a configuration $\bar{C} \in \mathcal{L}_t$. In this case, for each $C' \in \mathcal{R}_t$, either the edge (\bar{C}, C') is already assigned some demand (call such C' *blocked*) or C' has already been assigned $n_t(C')$ demands of type t . But there are at most D blocked C' 's, since the configuration \bar{C} has at most D slots to receive demands. Thus the total number of type- t demands assigned is at least

$$\sum_{C' \in \mathcal{R}_t \setminus \text{blocked}} n_t(C') \geq \sum_{C' \in \mathcal{R}_t} n_t(C') - D \cdot \max_{C' \in \text{blocked}} n_t(C') \geq \sum_{C' \in \mathcal{R}_t} n_t(C') - D \cdot D_t.$$

Moreover, exchanging the roles of \mathcal{L}_t and \mathcal{R}_t in the argument from (11) we get that $\sum_{C' \in \mathcal{R}_t} n_t(C') \geq n_t - \frac{D_t}{2}$, and thus at least $n_t - D \cdot D_t - \frac{D_t}{2}$ demands of type t are assigned by Step 4. Thus at most $O(D \cdot D_t)$ demands (hence total size $O(D)$) of this type remain unassigned and reach Step 5. This a total size of $O(DT)$, over all demand types, that reach the latter step, as desired. \square

We summarize the main results of this section and the previous with the next theorem: By approximately solving (LP_{mach}) ([Lemma 4](#)), rounding the solution ([Lemma 5](#)), and using

the above algorithm to obtain an assignment of demands to edges ([Theorem 6](#)), we obtain the following.

Theorem 7. *Consider an instance of OFFMINFAILOVER that has most T demands types and where each configuration has at most D demands. Then there is a poly-time algorithm that finds a feasible solution that uses at most $\text{LP}_{mach} + O(DT)$ machines.*

To use this procedure for obtaining our main result, [Theorem 4](#), we need to modify the input instance to make D and T small enough, which is the goal of the next section.

4.3 Reducing the number of types and demands in a configuration

Given an arbitrary set of demands J , we will convert this into another set of demands \tilde{J} with small parameters T and D such that: 1) The optimal LP value for J and \tilde{J} are similar; 2) We can convert an assignment of the demands \tilde{J} into an assignment of the original demands J without using many extra machines.

Given a parameter $\varepsilon \in (0, 1)$, the instance \tilde{J} is constructed as follows. First, partition $J = S \cup M$ into small ($s_j < \varepsilon^2$) and medium demands ($s_j \geq \varepsilon^2$), respectively. To reduce the number of types of medium demands we apply *linear grouping*, a transformation used in the context of Bin Packing [10]: Let $n := |M|$. Partition M into $\frac{1}{\varepsilon^3}$ groups, consisting of the $\varepsilon^3 n$ largest demands, the next $\varepsilon^3 n$ largest demands, and so on. Note every group has size exactly $\varepsilon^3 n$ except possibly the last group, corresponding to the smallest demands. At this point, let L be the first group, corresponding to the largest demands. We can now partition $J = S \cup (M \setminus L) \cup L$. For the remaining groups of demands in $M \setminus L$, let \tilde{M} denote the modified set of demands, where we round the size of each demand in $M \setminus L$ up to the largest size in its group. (The demands L will not be part of the final instance \tilde{J} and have to be handled separately.)

For the small demands S , we want to both reduce the number of types but also ensure that none of them are too small (in order to limit the max number of demands in a configuration). Simply rounding all of them to the threshold ε^2 may increase the LP value too much, and rounding down to 0 is effectively ignoring these demands, which makes it difficult to produce an assignment for the original instance J . So instead the idea is to group the small items S into “blocks” of size exactly ε . Since no set of small items may add to exactly this size, we actually just create the appropriate amount of “blocks”: let \tilde{S} be a collection of $\lceil \frac{1}{\varepsilon} \sum_{s \in S} s \rceil$ demands of size ε .

Then the transformed instance is given by the modified medium and small items, that is, $\tilde{J} = \tilde{S} \cup \tilde{M}$. The parameters T and D are indeed controlled: there are $T = \frac{1}{\varepsilon^3} + 1 = O(\frac{1}{\varepsilon^3})$ types, and since all demands have size at least ε^2 the max number of demands in a configuration is $D = \frac{1}{\varepsilon^2}$. (Also note that the large demands L , which are treated separately, have sizes at least ε^2 and there are at most $\varepsilon^3 |M|$ of them.) The next lemma states that our modifications did not increase LP_{mach} by much. For any set of demands J , we let $\text{LP}_{mach}(J)$ be the optimal LP value for this set of demands (recall that this is well-defined regardless of the partition of demands into types [Appendix E.1](#)).

Lemma 7. *For any set of demands J and $\varepsilon \in (0, 1)$, let the modified set of demands $\tilde{J} = \tilde{S} \cup \tilde{M}$ be defined as above. Then*

$$\text{LP}_{mach}(\tilde{J}) \leq (1 + O(\varepsilon))\text{LP}_{mach}(J) + O(1).$$

Proof. We first claim that $\text{LP}_{mach}(S \cup \tilde{M}) \leq \text{LP}_{mach}(J)$. To see this, recall that \tilde{M} is obtained from M by excluding L (the group of the largest $\varepsilon^3 n$ demands) and rounding up all remaining groups. Thus, for every group G in M except L , there exists a next-larger group G' in M with

at least as many demands and such that every demand in G' has size greater than or equal to every demand in G . This still holds after rounding up all demand sizes in G to largest demand in G , which is how we obtain \tilde{M} . Thus, we can monotonely match every item in $S \cup \tilde{M}$ to items in $S \cup M = J$, and with this it is not hard to see that $\text{LP}_{mach}(S \cup \tilde{M}) \leq \text{LP}_{mach}(J)$.

Given this inequality, it suffices to show $\text{LP}_{mach}(\tilde{S} \cup \tilde{M}) \leq (1 + O(\varepsilon))\text{LP}_{mach}(S \cup \tilde{M}) + O(1)$. We will use an optimal solution for $\text{LP}_{mach}(S \cup \tilde{M})$ to construct a feasible solution to $\text{LP}_{mach}(\tilde{S} \cup \tilde{M})$ without increasing the objective value by much. It is convenient to consider the type partition of $S \cup \tilde{M}$ where every small job is its own type, and there is one type for each group in \tilde{M} . Similarly, for $\tilde{S} \cup \tilde{M}$, we consider the type partition with one demand type for all demands in \tilde{S} and one for each group in \tilde{M} .

To achieve this, we map each configuration C of demands $S \cup \tilde{M}$ to a configuration C' using demands $\tilde{S} \cup \tilde{M}$ as follows: Starting from C , we keep all of its \tilde{M} demands. For the S (small) demands, let k be an integer such that the total size of all small demands in the configuration is in the interval $[k \cdot \varepsilon, (k + 1) \cdot \varepsilon)$. We arbitrarily remove small demands from this configuration until the total size of the remaining small demands lies in the interval $[k \cdot \varepsilon - \varepsilon^2, k \cdot \varepsilon]$. In doing so, we remove at most $\varepsilon + \varepsilon^2$ units of small demands (using the fact that every small demand has size at most ε^2 .) Then we replace the S demands in this configuration C with $\max(0, k - 1)$ blocks of small demands (i.e. demands in \tilde{S}), obtaining a configuration C' of demands $\tilde{S} \cup \tilde{M}$.

Proposition 1. *The configurations C' created by the above procedure are valid configuration.*

Proof. Let C be the input configuration of demands $S \cup \tilde{M}$ and C' the resulting configuration of demands $\tilde{S} \cup \tilde{M}$ (and throughout, let k be the one used in the transformation). We need to show that $\sum_{s \in C'} s \leq 1$ and $\sum_{s \in C'} s + \max_{s \in C'} s \leq B$. For the first inequality, both C and C' have the same \tilde{M} demands and the total size of S demands in C is at least $k \cdot \varepsilon$ and the total size of \tilde{S} demands in C is $\max\{0, (k - 1) \cdot \varepsilon\}$; thus $\sum_{s \in C'} s \leq \sum_{s \in C} s \leq 1$, the last step following from the validity of C .

It remains to show the second inequality. There are two cases to consider:

1. If $k \leq 1$ then C' contains no blocks of small demands but only the \tilde{M} demands that are also present in C . Then we have

$$\sum_{s \in C'} s + \max_{s \in C'} s \leq \sum_{s \in C} s + \max_{s \in C} s \leq B, \quad (12)$$

as desired (the last inequality following from the validity of C).

2. Otherwise $k \geq 2$. In this case, C' contains $k - 1 \geq 1$ blocks and the total size of \tilde{S} demands in C' is $(k - 1) \cdot \varepsilon$, while the total size of S demands in C is at least $k \cdot \varepsilon$. Since both configurations have the same \tilde{M} demands, this gives $\sum_{s \in C'} s \leq \sum_{s \in C} s - \varepsilon$. Further, because \tilde{S} demand has size exactly ε , we have

$$\max_{s \in C'} s \leq \max \left\{ \varepsilon, \max_{s \in C' \cap \tilde{M}} s \right\} = \max \left\{ \varepsilon, \max_{s \in C \cap \tilde{M}} s \right\} \leq \varepsilon + \max_{s \in C} s.$$

Combining both bounds gives the desired inequality (as in (12)).

Thus C' is a valid configuration, concluding the proof of the proposition. \square

Finally, we construct a feasible setting of the x -variables for $(\text{LP}_{mach}(\tilde{S} \cup \tilde{M}))$ with objective value at most $(1 + O(\varepsilon))\text{LP}_{mach}(S \cup \tilde{M}) + O(1)$, completing the proof. Let x^* be an optimal solution for $(\text{LP}_{mach}(S \cup \tilde{M}))$. Let \mathcal{C} be the collection of all possible configurations for demands

$S \cup \tilde{M}$ and \mathcal{C}' the collection of all possible configurations for demands $\tilde{S} \cup \tilde{M}$. For every $C' \in \mathcal{C}'$, we define $\bar{x}_{C'} = \sum_{C \in \mathcal{C} | C \rightarrow C'} x_C^*$, where $C \rightarrow C'$ denotes the event that the above procedure maps configuration C to C' . In words, we map the configurations chosen by x^* in C to configurations in C' . At this point, the objective value of \bar{x} is $\sum_{C' \in \mathcal{C}'} \bar{x}_{C'} = \text{LP}_{mach}(S \cup \tilde{M})$. It remains to modify \bar{x} so that it is feasible.

Note that $\text{LP}_{mach}(\tilde{S} \cup \tilde{M})$ and $\text{LP}_{mach}(S \cup \tilde{M})$ have the same constraints for the demands in \tilde{M} , and the mapping $C \rightarrow C'$ preserves the number of each type of medium demand. Thus, \bar{x} satisfies all medium demand constraints. It remains to satisfy the constraint for the demands in \tilde{S} .

This constraint says that the total number of demands from \tilde{S} in the configurations picked by the solution has to be at least twice the number of demands in \tilde{S} . Since every demand in \tilde{S} has exactly the same size ε , this constraint can be equivalently written in terms of sizes as follows (let $size(A) := \sum_{s \in A} s$ for $A \subseteq \mathbb{R}_+$):

$$\frac{1}{\varepsilon} \sum_{C' \in \mathcal{C}'} x_{C'} \cdot size(C' \cap \tilde{S}) \geq 2|\tilde{S}|. \quad (13)$$

We modify \bar{x} to satisfy this.

We know that since x^* is feasible for $(\text{LP}_{mach}(S \cup \tilde{M}))$, the total size of S demands it picks up is at least twice $size(S)$, that is,

$$\sum_{C \in \mathcal{C}} x_C^* \cdot size(C \cap S) \geq 2size(S). \quad (14)$$

Also, notice that for every mapped configurations $C \rightarrow C'$ we have $size(C \cap S) \leq size(C' \cap \tilde{S}) + 2\varepsilon$ (since $size(C \cap S) \leq (k+1) \cdot \varepsilon$ and $size(C' \cap \tilde{S}) \geq (k-1) \cdot \varepsilon$ for some k). Then using the definition of \bar{x} we get

$$\begin{aligned} \sum_{C' \in \mathcal{C}'} \bar{x}_{C'} \cdot size(C' \cap \tilde{S}) &= \sum_{C' \in \mathcal{C}'} \sum_{C \in \mathcal{C} | C \rightarrow C'} x_C^* \cdot size(C' \cap \tilde{S}) \geq \sum_{C \in \mathcal{C}} x_C^* \cdot size(C \cap S) - 2\varepsilon \sum_{C \in \mathcal{C}} x_C^* \\ &\geq 2size(S) - 2\varepsilon \text{LP}_{mach}(S \cup \tilde{M}), \end{aligned}$$

where the last inequality uses (14). Thus, it suffices to increase the \tilde{S} size of the demands picked up by \bar{x} by $2\varepsilon \cdot (1 + \text{LP}_{mach}(S \cup \tilde{M}))$ for it to satisfy (13). Dividing through by ε and using $|\tilde{S}| = \lceil \frac{1}{\varepsilon} \sum_{s \in \tilde{S}} s \rceil \leq \frac{size(\tilde{S})}{\varepsilon} + 1$ we get

$$\frac{1}{\varepsilon} \sum_{C' \in \mathcal{C}'} \bar{x}_{C'} \cdot size(C' \cap \tilde{S}) \geq 2|\tilde{S}| - 2 - 2\text{LP}_{mach}(S \cup \tilde{M}).$$

For \bar{x} to satisfy (13) we just need to increase it so it covers $\lceil 2 + 2\text{LP}_{mach}(S \cup \tilde{M}) \rceil$ extra demands of (type) \tilde{S} . For that, use [Lemma 6](#) to assign these additional demands to at most $O(\varepsilon) \cdot (1 + \text{LP}_{mach}(S \cup \tilde{M})) + O(1)$ many machines, and add to \bar{x} the configurations of these machines. Now \bar{x} satisfies (13), and has objective value at most $(1 + O(\varepsilon))\text{LP}_{mach}(S \cup \tilde{M}) + O(1)$. \square

4.4 Putting it all together

We finally obtain the complete algorithm for assigning the original demands J . At a high-level, we schedule the large demands L in edges by themselves, use the configuration LP plus rounding and realization of the configurations into an edge assignment ([Theorem 7](#)) for the instance \tilde{J} with modified small and medium demands to create a template, and finally replace them by the original small and (non-large) medium items S and $M \setminus L$. The precise algorithm is the following:

OffMinFailoverAlg: Given a collection J of demands, failover capacity $B \geq 1$, and parameter $\varepsilon \in (0, 1)$:

- 1: Construct the sets of small, medium, and large demands $J = S \cup M \cup L$ as well as the blocks of small demands and grouped medium demands \tilde{S} and \tilde{M} , respectively as in [Section 4.3](#).
 - 2: For each large demand, open a new edge (two machines) for that demand and assign it there.
 - 3: For the blocks and grouped medium demands $\tilde{S} \cup \tilde{M}$, there is a total of $T = O(\frac{1}{\varepsilon^3})$ demand types and each configuration can have at most $D = O(\frac{1}{\varepsilon^2})$ demands. So run the algorithm guaranteed by [Theorem 7](#) to obtain an assignment of $\tilde{S} \cup \tilde{M}$ into at most $\text{LP}_{mach}(\tilde{S} \cup \tilde{M}) + O(\frac{1}{\varepsilon^5})$ machines. We open this many machines and use the above assignment as a template to actually assign S and $M \setminus L$.
 - 4: Recall that we only increased the size of each demand from $M \setminus L$ to \tilde{M} , so we can assign each demand in $M \setminus L$ in the place of its corresponding demand in \tilde{M} .
 - 5: For the demands in S , we consider them in arbitrary order. When considering a demand $s \in S$, if there exists a block in the template (i.e. relative to a demand from \tilde{S}) assignment with less than $\varepsilon - s$ units of small demands assigned there, then assign demand s in this block. If no such block exists, then we assign all remaining small demands using [Lemma 6](#).
-

Proof of [Theorem 4](#). It is immediate that OffMinFailoverAlg runs in polynomial time and assigns all demands. Further, this assignment satisfies the Nominal and Failover constraints: it puts each large demand on its own matching edge, our template assignment of $\tilde{S} \cup \tilde{M}$ is feasible by [Theorem 7](#), we only assign smaller demands in $S \cup M$ than in the template, and our assignment of the remaining small demands is feasible by [Lemma 6](#).

It remains to show that the number of machines used is at most $(1 + O(\varepsilon))\text{LP}_{mach}(J) + O(\frac{1}{\varepsilon^5})$. We account the machines for the large, medium, and small demands separately.

1. For the large demands, we open $O(|L|) = O(\varepsilon^3|M|)$ machines. Moreover, every demand in M has size at least ε^2 , so every feasible configuration for the demands J has at most $\frac{1}{\varepsilon^2}$ such medium demands; since $\text{LP}_{mach}(J)$ needs to pick enough configurations to cover twice the medium demands, we get that $2|M| \leq \text{LP}_{mach}(J) \cdot \frac{1}{\varepsilon^2}$. Thus, $O(\varepsilon) \cdot \text{LP}_{mach}(J)$ machines are opened for the large demands.
2. For the medium demands, our template assignment of $\tilde{S} \cup \tilde{M}$ opens $\text{LP}_{mach}(\tilde{S} \cup \tilde{M}) + O(\frac{1}{\varepsilon^5}) \leq (1 + O(\varepsilon))\text{LP}_{mach}(J) + O(\frac{1}{\varepsilon^5})$, using [Lemma 7](#). We assign all medium demands in these machines.
3. For the small demands, it suffices to bound the number of extra machines needed for the remaining small demands that do not fit in the blocks. Consider the first time that we consider a small demand that cannot be assigned to any block. It must be the case that each block is already assigned at least $(\varepsilon - \varepsilon^2)$ units of small demands. Recall that the number of blocks is $\lceil \frac{1}{\varepsilon} \sum_{j \in S} s_j \rceil$. Thus the total size of already assigned small demands is at least $(\varepsilon - \varepsilon^2) \lceil \frac{1}{\varepsilon} \sum_{s \in S} s \rceil \geq (1 - \varepsilon) \sum_{j \in S} s_j$. We conclude that the total size of remaining unassigned small demands is at most $\varepsilon \cdot \sum_{s \in S} s = O(\varepsilon) \cdot \text{LP}_{mach}(J)$ (again $\text{LP}_{mach}(J)$ needs to cover twice all demands in S and each unit of configuration picked by J can cover at most 1 unit of size of these demands); this requires $O(\varepsilon) \cdot \text{LP}_{mach}(J)$ machines by [Lemma 6](#).

In total, we see that the algorithm uses at most $(1 + O(\varepsilon))\text{LP}_{mach}(J) + O(\frac{1}{\varepsilon^5})$ machines as claimed. This concludes the proof of [Theorem 4](#). \square

5 Rate of Convergence of the Minimum Number of Machines

The goal of this section is to understand the minimum number of machines needed to assign T items drawn i.i.d. from distribution μ . Recall we denote this random variable by $\text{OPT}_{mach}(X_1, \dots, X_T)$. Our main result here ([Theorem 5](#), restated) is that in expectation, this random variable is approximately linear in T .

Theorem 5. *Let μ be a distribution supported on $[0, \min\{1, \frac{B}{2}\}]$. Then there exists a scalar $c(\mu)$ such that for every $T \in \mathbb{N}$, we have*

$$\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) \in T \cdot c(\mu) \pm O(T^{5/6}),$$

where X_1, \dots, X_T are i.i.d. samples from μ .

One should interpret the constant $c(\mu)$ as the average number of machines needed per demand as the number of demands goes to infinity. Thus, by dividing both sides of the theorem by T , we have a quantitative convergence for the expected average number of machines needed per demand for T demands, $\frac{1}{T} \mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T)$, to the limiting value $c(\mu)$.

The main idea to prove [Theorem 5](#) is to consider a deterministic proxy for $\text{OPT}_{mach}(X_1, \dots, X_T)$. To construct this proxy we follow the approach of [\[13\]](#) that proves a similar result for the Bin Packing problem: for any distribution μ supported on $[0, 1]$, we define its *quantile function* μ^{-1}

$$\mu^{-1}(p) := \inf\{x \in [0, 1] : \mu([0, x]) \geq p\}.$$

For example, if μ is a continuous distribution, then $\mu^{-1}(p)$ is the unique value x such that $\mu([0, x]) = p$. Then, let μ_T denote the instance that has T demands whose sizes are given by

$$\left\{ \mu^{-1}\left(\frac{0}{T}\right), \mu^{-1}\left(\frac{1}{T}\right), \dots, \mu^{-1}\left(\frac{T-1}{T}\right) \right\}.$$

Note that μ_T is a deterministic instance. Roughly, in the stochastic instance $\{X_1, \dots, X_T\}$, we “expect” one demand to fall into each quantile $[\mu^{-1}(\frac{k}{T}), \mu^{-1}(\frac{k+1}{T})]$ for each $k = 0, \dots, T-1$. Thus, our deterministic proxy for $\text{OPT}_{mach}(X_1, \dots, X_T)$ is $\text{OPT}_{mach}(\mu_T)$.

Keeping in mind our interpretation for the constant $c(\mu)$ from before (the average number of machines needed per demand) and our deterministic proxy, we take $c(\mu) := \limsup_{n \rightarrow \infty} \frac{1}{n} \text{OPT}(\mu_n)$. We will show that this choice of $c(\mu)$ has the desired property.

There are two main steps to prove [Theorem 5](#). We first show that $\text{OPT}_{mach}(\mu_T)$ is a good proxy for $\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T)$. The proof of the next lemma relies on another Rhee-Talgrand-like monotone matching argument, where we construct a matching between the X_t 's and $\mu^{-1}(\frac{t}{T})$'s such that few demands are left unmatched.

Lemma 8. *For every $T \in \mathbb{N}$, we have*

$$\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) \in \text{OPT}_{mach}(\mu_T) \pm O(\sqrt{T}),$$

where X_1, \dots, X_T are i.i.d. samples from μ .

Second, we show that $\text{OPT}_{mach}(\mu_T)$ has the desired approximate linearity property. This relies on relating $\text{OPT}_{mach}(\mu_T)$ to its LP relaxation, $(\text{LP}_{mach}(\mu_T))$, whose optimal value is approximately linear in T .

Lemma 9. *For every $T \in \mathbb{N}$, we have $\text{OPT}_{mach}(\mu_T) \in T \cdot c(\mu) \pm O(T^{5/6})$.*

[Theorem 5](#) follows immediately from the above two lemmas, which we prove in the subsequent sections.

5.1 Proof of Lemma 8: $\text{OPT}_{mach}(\mu_T)$ is a good proxy

For convenience, we let $\mu_T = \{s_0, \dots, s_{T-1}\}$, where $s_j = \mu^{-1}(j/T)$. There are two analogous directions to prove: $\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) \leq \text{OPT}_{mach}(\mu_T) + O(\sqrt{T})$, and $\text{OPT}_{mach}(\mu_T) \leq \mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) + O(\sqrt{T})$. For the former, we use the assignment of μ_T into $\text{OPT}_{mach}(\mu_T)$ devices as a template to assign X_1, \dots, X_T using only $O(\sqrt{T})$ extra devices. To do so, we show that there exists a large *monotone matching* from the X_i 's to the s_j 's such that if X_i is matched to s_j , then $X_i \leq s_j$. We can bound the number of unmatched X_i 's using a quantitative version of Hall's theorem (Theorem 1.3.1 of [17]):

Theorem 8. *Let $G = (L \cup R, E)$ be a bipartite graph. For any subset $U \subseteq L$, we define its deficiency by $def(U) := |U| - |N(U)|$, where $N(U) \subseteq R$ is the set of neighbors of U . Then there exists a matching in G that leaves at most $\max_{U \subseteq L} def(U)$ vertices of L unmatched.*

For all matched X_i 's, we can assign them to the same position as their matched s_j -counterpart using $\text{OPT}_{mach}(\mu_T)$ machines as a template. For the unmatched ones, we assign them each to their own disjoint edge (opening 2 extra devices). To complete the proof, we need to show that in expectation, few of the X_i 's are unmatched. To do so, we use the Dvoretzky-Kiefer-Wolfowitz Inequality (Theorem 9 in Appendix A) to quantify the deviation of the empirical quantiles of the X_i 's with the "true" quantiles $\mu^{-1}(j/T)$. We now proceed formally.

Proposition 2. *We have $\mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) \leq \text{OPT}_{mach}(\mu_T) + O(\sqrt{T})$.*

Proof. Consider the (random) bipartite graph G with T vertices on each side such that the left side vertices correspond to the X_i 's and the right to the s_j 's. We have an edge (X_i, s_j) exactly when $X_i \leq s_j$.

Note that the maximum deficiency subset of the X_i 's (as defined in Theorem 8) must correspond to the random subset of all X_i 's that are strictly larger than some s_{j-1} for some $j = 1, \dots, T$. The deficiency of the j th such subset is $def(j) := \#\{X_i \text{'s strictly larger than } s_{j-1}\} - (T - j)$. Thus, a maximum matching in G leaves at most $\max_j def(j)$ of the X_i 's unmatched by Theorem 8. Fix some such maximum matching. We use it to assign the X_i 's as follows:

1. Open $\text{OPT}_{mach}(\mu_T)$ many devices. Consider the tentative assignment of μ_T to these devices.
2. For each matched X_i , we assign it to the pair of devices that its matched s_j is tentatively assigned to.
3. For each unmatched X_i , we open two more devices and assign X_i to the edge between them.

This is a feasible assignment of all X_i 's (because we assign each matched X_i to a slot for a larger s_j and each X_i fits on an edge by itself) using at most $\text{OPT}_{mach}(\mu_T) + 2 \cdot \max_j def(j)$ devices. It remains to show $\mathbb{E} \max_j def(j) = O(\sqrt{T})$. We re-write $def(j)$ using Claim 4 from Appendix B in terms of the tails of μ :

$$\begin{aligned} def(j) &= \#\{X_i \text{'s strictly larger than } s_{j-1}\} - (T - j) \\ &\leq \#\{X_i \text{'s strictly larger than } s_{j-1}\} - T \cdot \mu((s_{j-1}, 1]) + 1. \end{aligned}$$

Moreover, using the Dvoretzky-Kiefer-Wolfowitz Inequality (Theorem 9) we have for any $\lambda > \frac{1}{T}$

$$\begin{aligned} \Pr\left(\frac{1}{T} \max_j def(j) \geq \lambda\right) &\leq \Pr\left(\max_{v \in [0, 1]} \left(\frac{1}{T} \#\{X_i \text{'s strictly bigger than } v\} - \mu((v, 1])\right) \geq \lambda - \frac{1}{T}\right) \\ &\leq 2e^{-2T(\lambda - \frac{1}{T})^2}. \end{aligned}$$

Integrating the tail gives:

$$\begin{aligned}\mathbb{E} \frac{1}{T} \max_j def_j &\leq O\left(\frac{1}{\sqrt{T}}\right) + \int_{2/\sqrt{T}}^{\infty} \Pr\left(\frac{1}{T} \max_j def(j) \geq \lambda\right) d\lambda \\ &\leq O\left(\frac{1}{\sqrt{T}}\right) + \int_{2/\sqrt{T}}^{\infty} 2e^{-2T(\lambda - \frac{1}{T})^2} d\lambda = O\left(\frac{1}{\sqrt{T}}\right),\end{aligned}$$

where the last inequality can be seen, for example, by noticing that the integral is at most a constant times the mean of a folded normal distribution with standard deviation $1/\sqrt{2T}$, which is $O(1/\sqrt{T})$. Re-arranging gives $\mathbb{E} \max_j(def(j)) \leq O(\sqrt{T})$, as required. \square

The proof of the other direction is analogous.

Proposition 3. *We have $\text{OPT}_{mach}(\mu_T) \leq \mathbb{E} \text{OPT}_{mach}(X_1, \dots, X_T) + O(\sqrt{T})$.*

Proof. We again consider a random bipartite graph G on the same vertices but we switch the roles of the X_i 's and s_j 's. That is, now we have an edge (s_j, X_i) exactly when $s_j \leq X_i$. Analogously, the maximum deficiency subset of the s_j 's corresponds to some set $\{s_j, \dots, s_{T-1}\}$ for some $j = 0, \dots, T-1$ with deficiency $def(j) := T - j - \#\{X_i\text{'s at least } s_j\}$.

As before, we use the tentative assignment of demands X_1, \dots, X_T into $\text{OPT}_{mach}(X_1, \dots, X_T)$ devices and a maximum matching of G that leaves at most $\max_j def(j)$ of the s_j 's unmatched to assign μ_T . In particular, we open $\text{OPT}_{mach}(X_1, \dots, X_T)$ devices and assign each matched s_j to the slot of its matched X_i . For all remaining unmatched s_j 's, we assign them to disjoint edges. This gives a feasible assignment of the s_j 's into at most $\text{OPT}_{mach}(X_1, \dots, X_T) + 2 \cdot \max_j def(j)$ devices. It remains to show $\mathbb{E} \max_j def(j) = O(\sqrt{T})$.

Again using [Claim 4](#) we have

$$def(j) \leq T \cdot \mu([s_j, 1]) - \#\{X_i\text{'s at least } s_j\},$$

and so again using the DKW Inequality we get $\Pr(\frac{1}{T} \max_j def(j) \geq \lambda) \leq 2e^{-2T\lambda^2}$ for any $\lambda > 0$. An analogous calculation by integrating the tail gives $\mathbb{E} \max_j def(j) = O(\sqrt{T})$, as required. \square

To summarize, in both directions (from the X_i 's to s_j 's and the reverse), we can use a monotone matching and template assignment to find a good assignment of one type of demands from the other. Combining both propositions proves [Lemma 8](#).

5.2 Proof of [Lemma 9](#): Approximate linearity of $\text{OPT}_{mach}(\mu_T)$

Now we relate the optimum of the finite deterministic instances μ_T and the limit optimum $c(\mu) = \limsup_{n \rightarrow \infty} \frac{1}{n} \text{OPT}(\mu_n)$. Again we have two directions to prove. For the first (more difficult) direction, we relate $\text{OPT}_{mach}(\mu_T)$ with its LP relaxation, $\text{OPT}(\text{LP}_{mach}(\mu_T))$.

Proposition 4. *For all $T \in \mathbb{N}$ we have*

$$\text{OPT}_{mach}(\mu_T) \leq T \cdot c(\mu) + O(T^{5/6}).$$

Proof. It suffices to prove for $\varepsilon \in (0, 1)$:

$$\text{OPT}_{mach}(\mu_T) \leq (1 + O(\varepsilon)) T \cdot \frac{\text{OPT}_{mach}(\mu_{kT})}{kT} + O\left(\frac{1}{\varepsilon^5}\right) \quad (15)$$

for all integers $k \geq 1$. Then taking $\limsup_{k \rightarrow \infty}$ on both sides and noticing

$$\limsup_{k \rightarrow \infty} \frac{1}{kT} \text{OPT}_{mach}(\mu_{kT}) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \text{OPT}_{mach}(\mu_n) = c(\mu)$$

gives $\text{OPT}_{mach}(\mu_T) \leq (1 + O(\varepsilon))T \cdot c(\mu) + O(\frac{1}{\varepsilon^5})$ (recall that passing to a subsequence cannot increase a lim sup). Setting $\varepsilon = \frac{1}{T^{1/6}}$ to optimize the bound gives

$$\text{OPT}_{mach}(\mu_T) \leq T \cdot c(\mu) + O(T^{5/6})$$

as desired.

It remains to prove (15). Recall that $\text{OPT}(\text{LP}_{mach}(J))$ denotes the optimal value of the configuration LP (LP_{mach}) for the set of demands J . Then Theorem 4 allows us to bound the gap between $\text{OPT}_{mach}(\mu_T)$ and $\text{OPT}(\text{LP}_{mach}(\mu_T))$.

$$\text{OPT}_{mach}(\mu_T) \leq (1 + O(\varepsilon)) \text{OPT}(\text{LP}_{mach}(\mu_T)) + O\left(\frac{1}{\varepsilon^5}\right). \quad (16)$$

Moreover, notice that $\text{OPT}(\text{LP}_{mach})$ is linear with respect to duplicating items. That is, for every integer k we have

$$\text{OPT}(\text{LP}_{mach}(k \cdot \mu_T)) = k \cdot \text{OPT}(\text{LP}_{mach}(\mu_T)), \quad (17)$$

where $k \cdot \mu_T$ denotes the instance that has k copies of each item in μ_T .

It remains to relate $k \cdot \mu_T$ with μ_{kT} . Because the inverse CDF function μ^{-1} is non-decreasing, we can relate these two sets of demand as follows.

$$\begin{aligned} k \cdot \mu_T &= \left\{ \underbrace{\mu^{-1}\left(\frac{0}{T}\right), \dots, \mu^{-1}\left(\frac{0}{T}\right)}_{k \text{ times}}, \underbrace{\mu^{-1}\left(\frac{1}{T}\right), \dots, \mu^{-1}\left(\frac{1}{T}\right)}_{k \text{ times}}, \dots, \underbrace{\mu^{-1}\left(\frac{T-1}{T}\right), \dots, \mu^{-1}\left(\frac{T-1}{T}\right)}_{k \text{ times}} \right\} \\ \mu_{kT} &= \left\{ \underbrace{\mu^{-1}\left(\frac{0}{kT}\right), \dots, \mu^{-1}\left(\frac{k-1}{kT}\right)}_{\text{first } k \text{ items}}, \underbrace{\mu^{-1}\left(\frac{k}{kT}\right), \dots, \mu^{-1}\left(\frac{2k-1}{kT}\right)}_{\text{next } k \text{ items}}, \dots, \underbrace{\mu^{-1}\left(\frac{(T-1)k}{kT}\right), \dots, \mu^{-1}\left(\frac{Tk-1}{kT}\right)}_{\text{next } k \text{ items}} \right\} \end{aligned}$$

we see that the sizes in μ_{kT} dominate those in $k \cdot \mu_T$ (i.e., there is a perfect monotone matching from $k \cdot \mu_T$ to μ_{kT}). Then one can see that the optimal LP values for these instances satisfy the expected relationship $\text{OPT}(\text{LP}_{mach}(k \cdot \mu_T)) \leq \text{OPT}(\text{LP}_{mach}(\mu_{kT}))$. Together with (17) this gives

$$\text{OPT}(\text{LP}_{mach}(\mu_T)) = \frac{1}{k} \text{OPT}(\text{LP}_{mach}(k \cdot \mu_T)) \leq \frac{1}{k} \text{OPT}(\text{LP}_{mach}(\mu_{kT})).$$

Combining this bound with (16) completes the proof. \square

Finally, we need a converse to the above proposition

Proposition 5. *For every $T \in \mathbb{N}$ we have*

$$\text{OPT}_{mach}(\mu_T) \geq T \cdot c(\mu) - 2.$$

Proof. It suffices to show that for every n we have

$$\text{OPT}_{mach}(\mu_T) \geq T \cdot \frac{\text{OPT}_{mach}(\mu_n)}{n} - 2 - \frac{2T^2}{n}. \quad (18)$$

Taking the $\limsup_{n \rightarrow \infty}$ gives the desired result.

To prove (18), we fix n and write it as $n = kT + r$ for non-negative integers k, r with remainder $r < k$. We will upper bound $\text{OPT}_{mach}(\mu_n)$ as a function of $\text{OPT}_{mach}(\mu_T)$. To do so, we first construct the intermediate instance S obtained by increasing the size of the demands in μ_n as follows (recall that μ_n has items $\mu^{-1}(\frac{j}{kT+r})$ for $j = 0, \dots, kT + r - 1$):

- For every $j = 0, \dots, Tk - 1$, let \tilde{i} be such that $j \in [(\tilde{i} - 1)k, \tilde{i}k)$; then take the item $\mu^{-1}(\frac{j}{kT+r})$ of μ_n and increase its size to $\mu^{-1}(\frac{\tilde{i}k}{kT})$, and add the latter to S .
- For every $j = Tk, \dots, Tk + r - 1$, take the item $\mu^{-1}(\frac{j}{kT+r})$ of μ_n and increase its size to $\mu^{-1}(1)$, and add the latter to S .

We have only increased demand sizes from μ_n to S , so we have $\text{OPT}_{mach}(\mu_n) \leq \text{OPT}_{mach}(S)$. To further upper bound $\text{OPT}_{mach}(S)$, notice that S has the structure

$$S = \left\{ \underbrace{\mu^{-1}(\frac{k}{kT}), \dots, \mu^{-1}(\frac{k}{kT})}_{k \text{ times}}, \underbrace{\mu^{-1}(\frac{2k}{kT}), \dots, \mu^{-1}(\frac{2k}{kT})}_{k \text{ times}}, \dots, \underbrace{\mu^{-1}(\frac{Tk}{kT}), \dots, \mu^{-1}(\frac{Tk}{kT})}_{k \text{ times}}, \underbrace{\mu^{-1}(1), \dots, \mu^{-1}(1)}_{r \text{ times}} \right\},$$

which is exactly the union of k copies of the instance $\overline{\mu_T} := (\mu_T \cup \{\mu^{-1}(1)\}) \setminus \{\mu^{-1}(0)\}$ and the instance of “big” demands B that has r items of size $\mu^{-1}(1)$, i.e. $S = (k \cdot \overline{\mu_T}) \cup B$. Next, observe the subadditivity relation

$$\text{OPT}_{mach}(S) \leq k \cdot \text{OPT}_{mach}(\overline{\mu_T}) + \text{OPT}_{mach}(B),$$

since the optimal solutions of each of the instances $\overline{\mu_T}$ and B can be concatenated, giving a feasible solution for S with $k \cdot \text{OPT}_{mach}(\overline{\mu_T}) + \text{OPT}_{mach}(B)$ machines. Moreover, we claim that

$$\text{OPT}_{mach}(\overline{\mu_T}) \leq \text{OPT}_{mach}(\mu_T) + 2 \quad \text{and} \quad \text{OPT}_{mach}(B) \leq 2r :$$

The first inequality is because we can assign all demands $\mu_T \setminus \{\mu^{-1}(0)\}$ using at most $\text{OPT}_{mach}(\mu_T)$ machines and then assign the remaining demand $\mu^{-1}(1)$ using 2 extra machines; the second inequality is because a feasible solution for B is to assign each demand to 2 separate machines.

Putting all of these bounds together we obtain

$$\text{OPT}_{mach}(\mu_n) \leq \text{OPT}_{mach}(S) \leq k \text{OPT}_{mach}(\mu_T) + 2k + 2r.$$

Dividing though by n and using the facts $\frac{k}{n} \leq \frac{1}{T}$ and $\frac{r}{n} \leq \frac{T}{n}$ we get

$$\frac{1}{n} \text{OPT}_{mach}(\mu_n) \leq \frac{1}{T} \text{OPT}_{mach}(\mu_T) + \frac{2}{T} + \frac{2T}{n},$$

which is equivalent to the desired inequality (18). This concludes the proof. \square

Combining the above two propositions completes the proof of [Lemma 9](#). To summarize, for both propositions we needed the approximate linearity of $\text{OPT}_{mach}(\mu_n)$. In the former, we argued via LP relaxations, and in the latter by concatenating sub-instances.

References

- [1] AHMADIAN, S., AND FRIGGSTAD, Z. Further approximations for demand matching: Matroid constraints and minor-closed graphs. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland* (2017), I. Chatzigiannakis, P. Indyk, F. Kuhn, and A. Muscholl, Eds., vol. 80 of *LIPIcs*, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, pp. 55:1–55:13.
- [2] BALOGH, J., BÉKÉSI, J., DÓSA, G., EPSTEIN, L., AND LEVIN, A. A New and Improved Algorithm for Online Bin Packing. In *26th Annual European Symposium on Algorithms (ESA 2018)* (Dagstuhl, Germany, 2018), Y. Azar, H. Bast, and G. Herman, Eds., vol. 112 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, pp. 5:1–5:14.

- [3] BOUCHERON, S., LUGOSI, G., AND MASSART, P. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.
- [4] CHEKURI, C., AND KHANNA, S. A polynomial time approximation scheme for the multiple knapsack problem. *SIAM Journal on Computing* 35, 3 (2005), 713–728.
- [5] CSIRIK, J., JOHNSON, D. S., KENYON, C., ORLIN, J. B., SHOR, P. W., AND WEBER, R. R. On the sum-of-squares algorithm for bin packing. *Journal of the ACM (JACM)* 53, 1 (2006), 1–65.
- [6] GUPTA, A., AND KÖNEMANN, J. Approximation algorithms for network design: A survey. *Surveys in Operations Research and Management Science* 16, 1 (2011), 3–20.
- [7] GUPTA, A., AND MOLINARO, M. How the experts algorithm can help solve lps online. *Mathematics of Operations Research* 41, 4 (2016), 1404–1431.
- [8] GUPTA, V., AND RADOVANOVIĆ, A. Interior-point-based online stochastic bin packing. *Operations Research* 68, 5 (2020), 1474–1492.
- [9] HOBERG, R., AND ROTHVOSS, T. A logarithmic additive integrality gap for bin packing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (USA, 2017), SODA '17, Society for Industrial and Applied Mathematics*, p. 2616–2625.
- [10] KARMARKAR, N., AND KARP, R. M. An efficient approximation scheme for the one-dimensional bin-packing problem. In *23rd Annual Symposium on Foundations of Computer Science (SFCS 1982)* (1982), pp. 312–320.
- [11] KORUPOLU, M., MEYERSON, A., RAJARAMAN, R., AND TAGIKU, B. Coupled and k-sided placements: Generalizing generalized assignment. *Math. Program.* 154, 1–2 (dec 2015), 493–514.
- [12] LAU, L.-C., RAVI, R., AND SINGH, M. *Iterative Methods in Combinatorial Optimization*, 1st ed. Cambridge University Press, USA, 2011.
- [13] LIU, S., AND LI, X. Online bin packing with known T. *arXiv preprint arXiv:2112.03200* (2021).
- [14] MASSART, P. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The Annals of Probability* 18, 3 (1990), 1269–1283.
- [15] MEHTA, A., SABERI, A., VAZIRANI, U., AND VAZIRANI, V. Adwords and generalized online matching. *J. ACM* 54, 5 (oct 2007), 22–es.
- [16] PLOTKIN, S. A., SHMOYS, D. B., AND TARDOS, É. Fast approximation algorithms for fractional packing and covering problems. *Math. Oper. Res.* 20, 2 (1995), 257–301.
- [17] PLUMMER, M. D., AND LOVÁSZ, L. *Matching theory*. Elsevier, 1986.
- [18] RHEE, W. T., AND TALAGRAND, M. Optimal bin packing with items of random sizes II. *SIAM Journal on Computing* 18, 1 (1989), 139–151.
- [19] RHEE, W. T., AND TALAGRAND, M. On-line bin packing of items of random sizes, II. *SIAM Journal on Computing* 22, 6 (1993), 1251–1256.

- [20] ROTHVOSS, T. The entropy rounding method in approximation algorithms. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms* (USA, 2012), SODA '12, Society for Industrial and Applied Mathematics, p. 356–372.
- [21] SHMOYS, D. B., AND TARDOS, É. An approximation algorithm for the generalized assignment problem. *Mathematical programming* 62, 1 (1993), 461–474.
- [22] ZHANG, C., KUMBHARE, A. G., MANOUSAKIS, I., ZHANG, D., MISRA, P. A., ASSIS, R., WOOLCOCK, K., MAHALINGAM, N., WARRIER, B., GAUTHIER, D., ET AL. Flex: High-availability datacenters with zero reserved power. In *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)* (2021), IEEE, pp. 319–332.

Appendix

A Concentration Inequalities

We also need a couple of concentration inequalities, starting with McDiarmid's Inequality (Theorem 6.2 of [3]).

Lemma 10 (McDiarmid's Inequality). *Let $g : \mathcal{Z}^n \rightarrow \mathbb{R}$ be a function with the bounded differences property, i.e. for every two vectors $z, z' \in \mathcal{Z}^n$ that only differ in 1 coordinate we have $|g(z) - g(z')| \leq M$. If Z_1, \dots, Z_n are independent random variables taking values in \mathcal{Z} , then for all $\alpha > 0$*

$$\Pr(|g(Z) - \mathbb{E}g(Z)| \geq \alpha) \leq 2e^{-\frac{2\alpha^2}{nM^2}},$$

The next classical inequality can be found for example in Theorem 2.8 of [3].

Lemma 11 (Chernoff's Inequality). *Let Z_1, \dots, Z_n be independent random variables in $[0, 1]$. Then for all $\lambda > 0$*

$$\Pr\left(\sum_{t \leq n} Z_t - \mathbb{E} \sum_{t \leq n} Z_t \geq \lambda\right) \leq e^{-\frac{2\lambda^2}{n}},$$

and the same holds for the lower tail, i.e. replacing " $\geq \lambda$ " for " $\leq -\lambda$ ".

We also need the Dvoretzky-Kiefer-Wolfowitz inequality that bounds the rate of uniform convergence of the empirical cdf to the true cdf. To state it, given a distribution μ over the reals, let $\text{cdf}(x) = \Pr_{X \sim \mu}(X \leq x)$ denote its cdf, and given i.i.d. samples $X_1, \dots, X_n \sim \mu$ let $\text{cdf}_n(x) = \frac{1}{n} \sum_{i \leq n} \mathbf{1}(X_i \leq x)$ be the empirical cdf. The following version of the DKW inequality is Corollary 1 of [14].

Theorem 9 (DKW Inequality). *For any distribution μ , any number of samples n , and all $\lambda > 0$,*

$$\Pr\left(\max_{x \in \mathbb{R}} |\text{cdf}_n(x) - \text{cdf}(x)| \geq \lambda\right) \leq 2e^{-2n\lambda^2}.$$

B Auxiliary Results

Claim 4. *Consider any probability measure μ over $[0, 1]$. Letting $s_j = \mu^{-1}(j/T)$, we have for $j = 0, \dots, T - 1$:*

- $T - j \geq T \cdot \mu((s_{j-1}, 1]) - 1$
- $T - j \leq T \cdot \mu([s_j, 1])$.

Proof. For the first item, by definition of s_{j-1} we have (using continuity of measures w.r.t. decreasing sets) $\mu([0, s_{j-1}]) \geq \frac{j-1}{T}$, which reorganizing gives $j \leq T \cdot \mu([0, s_{j-1}]) + 1$; this implies $T - j \geq T(1 - \mu([0, s_{j-1}])) - 1$, which is the desired bound.

For the second item, we have (using continuity of measures for increasing sets) $\mu([0, s_j]) \leq \frac{j}{T}$, which gives $T - j \leq T(1 - \mu([0, s_j]))$, which is exactly what we need. \square

Lemma 12. *Consider non-negative constants c, d and $\alpha \in (0, 1)$. If $m, t > 0$ satisfy $m \geq ct - (dt)^\alpha$, then*

$$t \leq \max \left\{ \frac{1}{c} \left(m + \left(\frac{m(d+1)}{c} \right)^\alpha \right), \left(\frac{(d+1)d^\alpha}{c} \right)^{\frac{1}{1-\alpha}} \right\}. \quad (19)$$

Proof. Assume $t^{1-\alpha} > \frac{(d+1)d^\alpha}{c}$, otherwise we are done by the second term in the max. Multiplying through by t^α we obtain $t > \frac{(d+1)(dt)^\alpha}{c}$. Rewriting the left-hand-side as $t = \left(\frac{(d+1)c}{c} - d\right) \cdot t$ and reorganizing gives

$$\frac{d+1}{c} \left(ct - (dt)^\alpha \right) > dt.$$

Then the first term in the max in (19) multiplied by c is

$$m + \left(\frac{m(d+1)}{c} \right)^\alpha \stackrel{\star}{\geq} ct - (dt)^\alpha + \left(\frac{d+1}{c} \left(ct - (dt)^\alpha \right) \right)^\alpha \geq ct - (dt)^\alpha + (dt)^\alpha = ct,$$

where the inequality \star follows from the assumption $m \geq ct - (dt)^\alpha$. Dividing through by c proves (19) (using the first term in the max). \square

C Additional Results for Online Worst-Case

C.1 Upper bound

In this section, we show an upper bound of $\frac{1}{2}$ for deterministic online algorithms for the FAILOVER problem in the worst-case, proving that the algorithm we design in Section 2 is essentially tight.

Theorem 10. *No deterministic online algorithm can obtain competitive ratio better than $\frac{1}{2}$ for the FAILOVER problem in the worst-case model, even when $B = \infty$.*

Proof. Consider an instance with 4 devices and let $\varepsilon > 0$ be a parameter. Then, the demands arrive as follows:

- The first 2 demands have size ε .
- If the first 2 demands are placed on the same edge, then there are 2 more demands of size $1 - \varepsilon$.
- If the first 2 demands are not placed on the same edge, then there is one more demand of size 1.

Case 1: the first 2 demands go on the same edge. In this case, we can only fulfill one of the demands of size $1 - \varepsilon$ that can be placed on the two devices with no existing load, giving the algorithm total value of at most $1 + \varepsilon$. OPT can place all 4 demands, e.g., by placing the ε and $1 - \varepsilon$ demands alternatively on edges of a fixed 4-cycle. This gives optimal value 2. So, the competitive ratio of the algorithm is $\frac{1+\varepsilon}{2}$.

Case 2: either the first 2 demands go on disjoint edges or on edges that share an endpoint. In this scenario, the algorithm cannot place the demand of size 1, while OPT can place it by putting the two demands of size ε on a single edge. The competitive ratio in this case is $\frac{2\varepsilon}{1+2\varepsilon}$.

Taking $\varepsilon \rightarrow 0$ gives the desired result. \square

C.2 Algorithm for small demands

We assume there is $L \in \mathbb{N}$ such that all demands have size $\leq \frac{1}{L}$. We will show that there exists an online algorithm with competitive ratio that goes to 1 as the size of the largest demand goes to 0. In particular, the claim is the following:

Theorem 11. *If for some $L \in \mathbb{N}$ that is a square all demands have size at most $\frac{1}{L}$, there is an online algorithm for FAILOVER in the worst-case model that has competitive ratio at least $\left(1 - \min\left\{\frac{3}{\sqrt{L}}, \frac{m}{L}\right\}\right)$.*

C.2.1 Algorithm

The algorithm opens cliques of machines and schedules the demands on their edges using first-fit. In particular, for each demand, the algorithm first considers all edges of the first clique in order before continuing on the edges of the second clique (if one exists) and so on.

Specifically, the algorithm is the following:

1. If $m < 3\sqrt{L}$, open a single clique containing all m machines. Otherwise, when a new clique is needed, open a clique of size \sqrt{L} ; if only $\sqrt{L'} < \sqrt{L}$ vertices remain, then open a $\sqrt{L'}$ -clique.
2. When a demand arrives, assign it to an edge of an opened clique using first-fit making sure that for each clique of size m' the total load on each edge is at most

$$\alpha_{m'} = \min\left\{\frac{B}{m'}, \frac{1}{m' - 1}\right\}$$

If needed, open a new clique if possible.

3. If the demand cannot be scheduled, then stop.

C.2.2 Analysis

Given the capacities $\alpha_{m'}$ on each edge of a m' -clique, it is easy to check that the algorithm creates a valid placement that satisfies both Nominal and Failover capacities.

Assume that the algorithm was not able to place everything (otherwise it is OPT). Upon termination, there is a demand of size at most $\frac{1}{L}$ that could not be scheduled, therefore for each clique of size m' that the algorithm opened, all edges have load at least

$$\alpha_{m'} - \frac{1}{L}$$

As a result, the total load on all edges of a m' -clique is at least

$$\frac{1}{2} \cdot m' \cdot (m' - 1) \left(\alpha_{m'} - \frac{1}{L}\right) \tag{20}$$

We also know that OPT can achieve at most

$$\frac{1}{2} \cdot \min\{m, (m - 1)B\}. \tag{21}$$

We now consider the following cases:

Case 1: $m < 3\sqrt{L}$. In this case, the algorithm opens a single clique of size m and schedules load at least

$$\frac{1}{2} \cdot m \cdot (m-1) \left(\alpha_m - \frac{1}{L} \right) \quad (22)$$

Depending on the value of B and m , we have the following cases:

- If $m < (m-1)B$, then $\text{OPT} \leq \frac{m}{2}$, and $\alpha_m = \frac{1}{m-1}$. From (22), we have:

$$\text{Alg} \geq \frac{1}{2} \cdot m \cdot (m-1) \left(\frac{1}{m-1} - \frac{1}{L} \right) = \frac{m}{2} \left(1 - \frac{m-1}{L} \right) \geq \text{OPT} \left(1 - \frac{m}{L} \right).$$

- If $m \geq (m-1)B$, then $\text{OPT} \leq \frac{1}{2}(m-1)B$, and $\alpha_m = \frac{B}{m}$. From (22), we have:

$$\text{Alg} \geq \frac{1}{2} \cdot m \cdot (m-1) \left(\frac{B}{m} - \frac{1}{L} \right) = \frac{(m-1)B}{2} \left(1 - \frac{m}{BL} \right) \geq \text{OPT} \left(1 - \frac{m}{L} \right).$$

In both cases, $\text{Alg} \geq \text{OPT} \left(1 - \frac{m}{L} \right) \geq \text{OPT} \left(1 - \min \left\{ \frac{3}{\sqrt{L}}, \frac{m}{L} \right\} \right)$.

Case 2: $m \geq 3\sqrt{L}$. Let C denote the number of \sqrt{L} -cliques that the algorithm opens. By construction, $C \cdot \sqrt{L} + \sqrt{L'} = m$.

\sqrt{L} -cliques. From (20), the overall load that the algorithm scheduled successfully on these cliques is at least

$$C \cdot \frac{1}{2} \cdot \sqrt{L} \cdot (\sqrt{L} - 1) \left(\alpha_{\sqrt{L}} - \frac{1}{L} \right) \geq C \frac{\sqrt{L} \cdot (\sqrt{L} - 1)}{2} \left(\frac{1}{\sqrt{L}} - \frac{1}{L} \right) \geq \frac{1}{2} C \sqrt{L} \left(1 - \frac{2}{\sqrt{L}} \right). \quad (23)$$

$\sqrt{L'}$ -clique. Similarly, for the $\sqrt{L'}$ -clique, the total load that is scheduled on this clique is at least

$$\frac{1}{2} \cdot \sqrt{L'} \cdot (\sqrt{L'} - 1) \left(\alpha_{\sqrt{L'}} - \frac{1}{L} \right) \geq \frac{\sqrt{L'} \cdot (\sqrt{L'} - 1)}{2} \left(\frac{1}{\sqrt{L'}} - \frac{1}{L} \right) \geq \frac{1}{2} \sqrt{L'} \left(1 - \frac{\sqrt{L'}}{L} - \frac{1}{\sqrt{L'}} \right). \quad (24)$$

Combining the above, the algorithm gets value at least

$$\begin{aligned} \text{Alg} &\geq \frac{1}{2} C \sqrt{L} \left(1 - \frac{2}{\sqrt{L}} \right) + \frac{1}{2} \sqrt{L'} \left(1 - \frac{\sqrt{L'}}{L} - \frac{1}{\sqrt{L'}} \right) \\ &= \frac{1}{2} \left(m - 2 \frac{m}{\sqrt{L}} + 2 \frac{\sqrt{L'}}{\sqrt{L}} - \frac{L'}{L} - 1 \right) \\ &\geq \frac{1}{2} \left(m - 2 \frac{m}{\sqrt{L}} - 1 \right) \geq \frac{m}{2} \left(1 - \frac{3}{\sqrt{L}} \right) \\ &\geq \text{OPT} \left(1 - \frac{3}{\sqrt{L}} \right) \geq \text{OPT} \left(1 - \min \left\{ \frac{3}{\sqrt{L}}, \frac{m}{L} \right\} \right) \end{aligned}$$

where we used the fact that $C \cdot \sqrt{L} + \sqrt{L'} = m \Leftrightarrow C = \frac{m - \sqrt{L'}}{\sqrt{L}}$, $L' < L$, and $\frac{m}{\sqrt{L}} \geq 1$.

D Omitted Proofs from Section 3

D.1 Proof of Lemma 1

In order to upper bound the utilization that OPT can achieve, the first step is to bound the minimum number of machines that are required to schedule a set of demands (with high probability). This allows us to generate bounds for the number of demands that can be scheduled (with high probability) when the number of machines is fixed. Connecting this number of demands with their sizes produces the desired bounds for OPT's utilization.

In particular, recall that $\text{OPT}_{mach}(s_1, \dots, s_n)$ denotes the minimum number of devices needed to schedule all the demands s_1, \dots, s_n satisfying the Nominal and Failover constraints. Also recall that [Theorem 5](#) shows that there exists a scalar $c(\mu)$ such that for every T , we have

$$\mathbb{E} \text{OPT}_{mach}(S_1, \dots, S_T) \in T \cdot c(\mu) \pm O(T^{5/6}).$$

However, for proving [Lemma 1](#), we need the observation that this bound holds not only in expectation but with high probability (with a negligible additional loss). To see that, notice that the function $\text{OPT}_{mach}(s_1, \dots, s_n)$ has bounded differences: changing the size of any demand i from s_i to s'_i (both in the range $[0, \min\{1, \frac{B}{2}\}]$) can change the minimum number of machines required by at most 2, i.e.,

$$|\text{OPT}_{mach}(s_1, \dots, s_i, \dots, s_n) - \text{OPT}_{mach}(s_1, \dots, s'_i, \dots, s_n)| \leq 2,$$

since we can always schedule the demand i on an edge by itself (using 2 new machines) if needed. Then as a consequence of McDiarmid's Inequality ([Lemma 10](#)) with $\alpha = \lambda\sqrt{T}$ we directly obtain the following.

Corollary 1. *For every T we have*

$$\text{OPT}_{mach}(S_1, \dots, S_T) \in T \cdot c(\mu) \pm O(T^{5/6}) \pm \lambda\sqrt{T}$$

with probability at least $1 - 2e^{-\frac{\lambda^2}{2}}$.

Then by essentially inverting this bound, we can upper bound how many demands are scheduled by the optimal solution $\text{OPT}(S_1, \dots, S_n)$ for our original problem.

Lemma 13. *Consider the original problem FAILOVER with m machines. For every $\delta > 0$, with probability at least $1 - \delta$ the optimal solution to the instance $S_1, \dots, S_n \sim \mu$ schedules at most $\frac{m}{c(\mu)} + O(m^{5/6}) + \log^{3/2} \frac{1}{\delta}$ demands.*

Proof sketch. Using [Corollary 1](#) with $\lambda = \sqrt{2 \log \frac{2}{\delta}}$, one can see that there is a positive constant d such that with probability at least $1 - \delta$

$$\text{OPT}_{mach}(S_1, \dots, S_T) \geq T \cdot c(\mu) - d^{5/6} \cdot T^{5/6} \tag{25}$$

for every number of demands $T \geq \log^{3/2} \frac{1}{\delta}$.

A bit of algebra shows that there is a value $\bar{T} = \frac{m}{c(\mu)} + O(m^{5/6}) + \log^{3/2} \frac{1}{\delta}$ such that the right-hand side of (25) is strictly more than m (the term $+\log^{3/2} \frac{1}{\delta}$ being present just to ensure $\bar{T} \geq \log^{3/2} \frac{1}{\delta}$); in fact, taking

$$\bar{T} := \max \left\{ \frac{1}{c(\mu)} \left(m + \left(\frac{m(d+1)}{c(\mu)} \right)^{5/6} \right), \frac{(d+1) d^{5/6}}{c(\mu)}, \log^{3/2} \frac{1}{\delta} \right\} + 1$$

suffices, which can be verified by applying the contrapositive of [Lemma 12](#)). This means that with probability at least $1 - \delta$, the demands $S_1, \dots, S_{\bar{T}}$ cannot all be scheduled within m machines; in such scenarios the optimal solution then schedules at most $\bar{T} - 1 = \frac{m}{c(\mu)} + O(m^{5/6}) + \log^{3/2} \frac{1}{\delta}$ demands, as claimed. \square

We can now show that with high probability $\text{OPT}(S_1, \dots, S_n) \leq m \cdot \frac{\mathbb{E}S_0}{c(\mu)} + O(m^{5/6})$ and conclude the proof of [Lemma 1](#). Setting $\delta = \frac{1}{m^2}$ and letting $\bar{T} = \frac{m}{c(\mu)} + O(m^{5/6})$ be the above upper bound on the number of demands scheduled by $\text{OPT}(S_1, \dots, S_n)$, with probability $\geq 1 - \frac{1}{m^2}$ we have that $\text{OPT}(S_1, \dots, S_n) \leq \sum_{t \leq \bar{T}} S_t$. Moreover, employing the Chernoff bound ([Lemma 11](#)) with $\lambda = \sqrt{\bar{T} \log m}$, with probability at least $1 - \frac{1}{m^2}$ this sum can be upper bounded as

$$\sum_{t \leq \bar{T}} S_t \leq \bar{T} \cdot \mathbb{E}S_0 + \sqrt{\bar{T} \log m} \leq m \cdot \frac{\mathbb{E}S_0}{c(\mu)} + O(m^{5/6}).$$

Taking a union bound to combine the two previous bounds, with probability at least $1 - \frac{2}{m^2}$ we have that $\text{OPT}(S_1, \dots, S_n) \leq m \cdot \frac{\mathbb{E}S_0}{c(\mu)} + O(m^{5/6})$, as desired. This concludes the proof of the [Lemma 1](#).

D.2 Proof of [Lemma 3](#)

To prove [Lemma 3](#) it will suffice to show that with high probability [OneRound](#) consumes a quarter of the machines available, unless there are already few machines available.

Lemma 14. *There is a constant cst_5 such that whenever $\tilde{m} \geq 4cst_5 \cdot m^{5/6} + 8$, we have that $\text{OneRound}(\tilde{m})$ opens at least $\frac{\tilde{m}}{4}$ machines with probability at least $1 - \frac{\log(m/c(\mu))}{m^2}$.*

Proof. The main element of the proof is a converse to [Claim 1](#), that is, a lower bound on the number of phases performed by [OneRound](#). To argue that it has not ran out of machines on an initial phase, let UB_ℓ be the upper bound on the number of machines opened by the [OneRound](#) until phase ℓ given by [\(10\)](#), namely

$$UB_\ell = 2n_\ell \cdot c(\mu) + cst_4 \cdot n_\ell^{5/6} + cst_4 \cdot m^{5/6}$$

for a sufficiently large constant cst_4 . Here is the desired bound on the number of phases performed.

Claim 5. *Let \underline{k} be the largest integer ℓ such that $UB_\ell + 2m^{5/6} \leq \tilde{m}$. Then with probability at least $1 - \frac{\underline{k}}{m^2}$ the algorithm $\text{OneRound}(\tilde{m})$ performs at least \underline{k} phases.*

Proof. From [\(10\)](#) and the fact that cst_4 and m are at least a sufficiently large constant, we have with probability at least $1 - \frac{\underline{k}}{m^2}$

$$Open_{\underline{k}} + cst_1 \cdot \sqrt{n_{\underline{k}}} \log^{3/4} n_{\underline{k}} \leq UB_{\underline{k}}. \quad (26)$$

Under this event we see that [OneRound](#) starts phase \underline{k} , that is, it does not STOP in Step (b) in the beginning of this iteration, since at that point

$$\begin{aligned} & \# \text{ already open machines} + \overline{\text{OPT}}_{mach}(Y_1, \dots, Y_{n_k}) + cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k + 2m^{5/6} \\ &= Open_{\underline{k}-1} + M_k + cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k + 2m^{5/6} \\ &\leq Open_{\underline{k}} + cst_1 \cdot \sqrt{n_k} \log^{3/4} n_k + 2m^{5/6} \\ &\leq UB_{\underline{k}} + 2m^{5/6} \leq \tilde{m}, \end{aligned}$$

where the last inequality follows from the definition of \underline{k} . Moreover, under this event the algorithm also does not run out of machines (i.e. fails) on this phase \underline{k} , since again $Open_{\underline{k}} \leq UB_{\underline{k}} \leq \tilde{m}$. The claim then follows. \square

Using this claim plus the lower bound on the number of machines M_k open on a round k from [Claim 2](#), we have that with probability at least $1 - (\underline{k} + 1)/m^2$

$$\begin{aligned}
Open &\geq Open_{\underline{k}} \geq \sum_{k=k_0}^{\underline{k}} M_k \geq \sum_{k=k_0}^{\underline{k}} \left(n_k \cdot c(\mu) - 2cst_1 \cdot n_k^{5/6} \right) \\
&\geq \sum_{k=1}^{\underline{k}} n_k \cdot c(\mu) - \sum_{k=1}^{k_0} n_k \cdot c(\mu) - \sum_{k=1}^{\underline{k}} 2cst_1 \cdot n_k^{5/6} \\
&\geq (2n_{\underline{k}} - 1) c(\mu) - 2n_{k_0} - 5cst_1 \cdot n_{\underline{k}}^{5/6} \\
&\geq (2n_{\underline{k}} - 1) c(\mu) - 2m^{5/6} - \frac{5cst_1}{c(\mu)^{5/6}} \cdot m^{5/6}, \tag{27}
\end{aligned}$$

where the next-to-last equality follows from the definition $n_k = 2^k$, and the last inequality uses [Claim 2](#) and that the definition of \underline{k} implies $n_{\underline{k}} \leq \frac{\tilde{m}}{c(\mu)} \leq \frac{m}{c(\mu)}$.

To further lower bound this quantity, from the maximality of \underline{k} we have $UB_{\underline{k}+1} + 2m^{5/6} > \tilde{m}$, which expanding the definition of $UB_{\underline{k}+1}$ and again using $n_{\underline{k}} \leq \frac{m}{c(\mu)}$ gives

$$2n_{\underline{k}} \cdot c(\mu) > \frac{1}{2} \left(\tilde{m} - cst_4 \cdot n_{\underline{k}+1}^{5/6} - (cst_4 + 2)m^{5/6} \right) \geq \frac{\tilde{m}}{2} - \left(\frac{cst_4}{2^{1/6}c(\mu)^{5/6}} + \frac{cst_4}{2} + 1 \right) m^{5/6}.$$

Employing this on inequality (27) we get that with probability at least $1 - 2\underline{k}/m^2$

$$Open \geq \frac{\tilde{m}}{2} - \underbrace{\left(\frac{cst_4/(2^{1/6}) + 5cst_1}{c(\mu)^{5/6}} + \frac{cst_4}{2} + 3 \right)}_{=: cst_5} m^{5/6} - 2.$$

Under the assumption $\tilde{m} \geq 4cst_5 \cdot m^{5/6} + 8$ of [Lemma 14](#), we get $Open \geq \frac{\tilde{m}}{4}$ as desired. To finalize, again since $n_{\underline{k}} \leq \frac{m}{c(\mu)}$, we have $\underline{k} \leq \log \frac{m}{c(\mu)}$, and so this happens with probability at least $1 - \frac{\log(m/c(\mu))}{m^2}$. This concludes the proof of [Lemma 14](#). \square

We now turn to the proof of the main lemma.

Proof of [Lemma 3](#). We assume that m is at least a sufficiently large constant, otherwise the result directly holds (with an appropriate constant in the term $O(\frac{1}{m})$); in particular, we assume that $5cst_5 \cdot m^{5/6} \geq 4cst_5 \cdot m^{5/6} + 8$, which will be useful to clean up the bounds.

Let \tilde{m}_i denote the number of unopened machines right before the i -th call to `OneRound` made by the main algorithm (i.e., at this call \tilde{m}_i is the parameter passed to `OneRound`). Define the (bad) event E_i that in the beginning of round i we still have more than $5cst_5 \cdot m^{5/6}$ unopened machines but (unlike what is prescribed by [Lemma 14](#) above) we did not consume open at least a quarter of these machines, i.e.

$$E_i \equiv (\tilde{m}_i > 5cst_5 \cdot m^{5/6}) \text{ and } (\tilde{m}_{i+1} > \frac{3}{4}\tilde{m}_i).$$

We claim that when neither of the events $E_1, \dots, E_{\bar{r}}$ holds (for $\bar{r} := \frac{\log m}{\log 4/3}$), then the total number of machines opened by the main algorithm is at least $m - 5cst_5 \cdot m^{5/6}$, which is what we want. To see this claim, notice that in this situation there are two cases:

Case 1: There is an event E_i ($i \leq \bar{r}$) which does not hold because $\tilde{m}_i \leq 5cst_5 \cdot m^{5/6}$. But this means that at the beginning of round i the main algorithm has already opened $m - \tilde{m}_i \geq m - 5cst_5 \cdot m^{5/6}$ machines, and the claim holds.

Case 2: All the events $\{E_i\}_{i \leq \bar{r}}$ do not hold because $\tilde{m}_{i+1} \leq \frac{3}{4}\tilde{m}_i$ for all of them. But this means that in beginning of the last round \bar{r} there are

$$\tilde{m}_{\bar{r}} \leq \left(\frac{3}{4}\right)^{\bar{r}-1} \cdot \tilde{m}_1 = \left(\frac{3}{4}\right)^{\bar{r}-1} \cdot m = \frac{4}{3} \leq 5cst_5 \cdot m^{5/6}$$

unopened machines, and so the claim also holds.

So to prove [Lemma 3](#) it suffices to show that the probability that an event $\{E_i\}_{i \leq \bar{r}}$ holds is at most $O(\frac{1}{m})$. This probability is

$$\Pr\left(\bigvee_{i \leq \bar{r}} E_i\right) \leq \sum_{i \leq \bar{r}} \Pr(E_i). \quad (28)$$

To upper bound the right-hand side, we have

$$\Pr(E_i) = \Pr(\tilde{m}_{i+1} > \frac{3}{4}\tilde{m}_i \mid \tilde{m}_i > 5cst_5 \cdot m^{5/6}) \Pr(\tilde{m}_i > 5cst_5 \cdot m^{5/6}).$$

But conditioning on $\tilde{m}_i > 5cst_5 \cdot m^{5/6}$ (or more precisely, conditioning on the demands up to the beginning of round i so that this event holds) and applying [Lemma 14](#) (notice that even with this conditioning the items within round i are still sampled i.i.d. from μ) we have that the first term in the right-hand side is at most $\frac{\log(m/c(\mu))}{m^2}$; so this gives $\Pr(E_i) \leq \frac{\log(m/c(\mu))}{m^2}$. Employing this on (28) gives

$$\Pr\left(\bigvee_{i \leq \bar{r}} E_i\right) \leq \bar{r} \cdot \frac{\log(m/c(\mu))}{m^2} = O\left(\frac{\log^2 m}{m^2}\right) = O\left(\frac{1}{m}\right).$$

This concludes the proof of [Lemma 3](#). □

E Properties of (LP_{mach})

E.1 Consistency of LP_{mach} with respect to type partitions

In principle, the definition of (LP_{mach}) depends on how the demands are partitioned into types. However, we show that this is not actually the case.

Lemma 15. *The optimal value LP_{mach} is the same for every possible type partition of the demands.*

In particular, the optimal value LP_{mach} is well-defined. To prove the lemma, it suffices to show that merging two demand types (that can be merged) does not change the optimal value.

Lemma 16. *Consider an instance $S = (s_1, \dots, s_k)$ and a valid assignment of types $type : S \rightarrow \{0, 1, \dots, k\}$ (i.e. $s = s'$ whenever $type(s) = type(s')$) and such that types 0 and 1 have items of the same size (i.e. $s = s'$ whenever $type(s), type(s') \in \{0, 1\}$). Let \widetilde{type} be the type assignment that merges types 0 and 1, i.e. $\widetilde{type}(s) = 1$ for all s such that $type^{-1}(s) \in \{0, 1\}$ and $\widetilde{type}(s) = type(s)$ for the other s 's.*

Then the optimal value of LP_{mach} based on $type$ and \widetilde{type} are the same.

Proof. Let $n_j = |type^{-1}(S)|$ be the number of items of type j under the assignment $type$ and define \tilde{n}_j analogously w.r.t. \widetilde{type} ; notice $\tilde{n}_1 = n_0 + n_1$ and $\tilde{n}_j = n_j$ for all $j \geq 2$. Further, for any configuration C w.r.t. $type$, we define $n_t(C)$ to be the number of demands of type t in C . We define $\tilde{n}_t(\tilde{C})$ analogously for a configuration \tilde{C} w.r.t. \widetilde{type} . The LP values are then given by

$$\begin{aligned}
LP &:= \min \sum_C x_C & \widetilde{LP} &:= \min \sum_{\tilde{C}} x_{\tilde{C}} \\
st \sum_C n_t(C) \cdot x_C &\geq 2n_t \quad \forall t & st \sum_{\tilde{C}} \tilde{n}_t(\tilde{C}) \cdot x_{\tilde{C}} &\geq 2\tilde{n}_t \quad \forall t \geq 1 \\
x &\geq 0 & x &\geq 0,
\end{aligned}$$

where the configurations C and \tilde{C} are respectively in \mathbb{R}^{k+1} and \mathbb{R}^k .

The second LP is a relaxation of the first, since it follows by adding the first two inequalities ($t = 0$ and $t = 1$) of the first LP. Thus, $\widetilde{LP} \leq LP$.

Now we prove that $LP \leq \widetilde{LP}$. Consider an optimal solution \tilde{x} for the second LP. Given a valid configuration \tilde{C} for the second LP, define the configurations $C^0 = (\tilde{C}_1, 0, \tilde{C}_2, \dots, \tilde{C}_n)$ and $C^1 = (0, \tilde{C}_1, \tilde{C}_2, \dots, \tilde{C}_n)$ that respectively assign all the items of \widetilde{type} to *type* 0 and 1. Consider the solution x for the first LP given by

$$x_{C^0} = \frac{n_0}{n_0 + n_1} \tilde{x}_{\tilde{C}}, \quad x_{C^1} = \frac{n_1}{n_0 + n_1} \tilde{x}_{\tilde{C}} \quad \text{for all } \tilde{C}$$

and $x(C) = 0$ for all other configurations.

We claim that x is a feasible solution for the first LP with value \widetilde{LP} . For its value

$$\sum_C x_C = \sum_{\tilde{C}} (x_{C^0} + x_{C^1}) = \sum_{\tilde{C}} \tilde{x}_{\tilde{C}} = \widetilde{LP},$$

as claimed. For its feasibility, for any $t \in \{0, \dots, k\}$ we have

$$\sum_C n_t(C) \cdot x_C = \sum_{\tilde{C}} (n_t(C^0) \cdot x_{C^0} + n_t(C^1) \cdot x_{C^1}) = \sum_{\tilde{C}} \left(\frac{n_0}{n_0 + n_1} n_t(C^0) + \frac{n_1}{n_0 + n_1} n_t(C^1) \right) \cdot \tilde{x}_{\tilde{C}}.$$

When $t \in \{0, 1\}$, we see that $n_t(C^t) = \tilde{n}_1(\tilde{C})$ and the other term $n_t(C^{1-t})$ is zero. Hence

$$\sum_C n_t(C) \cdot x_C = \frac{n_t}{n_0 + n_1} \sum_{\tilde{C}} \tilde{n}_1(\tilde{C}) \cdot \tilde{x}_{\tilde{C}} \geq \frac{n_t}{n_0 + n_1} \cdot 2\tilde{n}_1 = 2n_t,$$

where the first inequality is from the feasibility of \tilde{x} . So x satisfies the constraints of the first LP when $t = 0, 1$. For the remaining constraints $t \geq 2$ we have that $n_t(C^0) = n_t(C^1) = \tilde{n}_t(\tilde{C})$, so

$$\sum_C n_t(C) \cdot x_C = \sum_{\tilde{C}} \tilde{n}_t(\tilde{C}) \cdot \tilde{x}_{\tilde{C}} \geq 2\tilde{n}_t = 2n_t,$$

which are then satisfied as well. This proves that x is feasible for the first LP.

We conclude, the optimal value of this LP is at most that of this solution x , which then gives $LP \leq \widetilde{LP}$ as desired. \square

E.2 Solving (LP_{mach})

In this section, we show how to efficiently solve LP_{mach} up to small additive error. We need the following theorem of Rothvoss (stated in simplified form) [20], which relies on the Plotkin-Shmoys-Tardos algorithm to solve implicit fractional covering problems [16].

Theorem 12. Let $\mathcal{S} \subset 2^{[n]}$ be a set family. Suppose that we can solve the following SUBPROBLEM: Given parameter $\epsilon \in (0, 1)$ and $y \in \mathbb{Q}_+^n$, output a set $S^* \in \mathcal{S}$ with $\sum_{i \in S^*} y_i \geq (1 - \epsilon) \cdot \max_{S \in \mathcal{S}} (\sum_{i \in S} y_i)$ in time $T(n, \epsilon)$.

Then for any $\delta \in (0, n/2]$, we can find a basic solution of the following LP:

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}} x_S \\ \text{s.t.} \quad & \sum_{S \in \mathcal{S}} 1_{j \in S} \cdot x_S \geq 1 \quad \forall j \in [n] \\ & x \geq 0 \end{aligned} \tag{LP}$$

with objective value at most $LP + \delta$ in time $\text{poly}(n, 1/\delta) \cdot T(n, \Omega(\delta/n))$.

SUBPROBLEM is an approximate dual separation problem for LP. To apply this theorem, we relate (LP_{mach}) to a “per-demand” configuration LP, which we can apply the theorem to.

Lemma 17. Given (LP_{mach}) defined on n demands (with some partition into demands types) and a parameter $\delta \in (0, n/2]$, we can efficiently find a basic solution of (LP_{mach}) with objective value at most $LP_{mach} + \delta$ in time $\text{poly}(n, 1/\delta)$.

Proof. We let C be the collection of feasible configurations with respect to that partition into demand types in (LP_{mach}) (i.e. C indexes all variables used by this LP.) We first define the natural “per-demand” configuration LP by taking each demand as its own type. Indexing the demands by $j \in [n]$, this LP is:

$$\begin{aligned} \min \quad & \sum_{C' \in C_J} x_{C'} \\ \text{s.t.} \quad & \sum_{C' \in C_J} 1_{j \in C'} \cdot x_{C'} \geq 2 \quad \forall j \in [n] \\ & x \geq 0, \end{aligned} \tag{LP}_J$$

where C_J is the collection of all configurations where each demand is its own type. By Lemma 15, we have $LP_{mach} = LP_J$. To apply the theorem, we re-scale the right hand side of LP_J by dividing by 2. Let LP'_J be the resulting LP. Note that the extreme points and optimal solutions of LP_J and LP'_J are also related by a multiplicative 2-factor.

We now show how to solve SUBPROBLEM for LP'_J . We are given parameter $\epsilon \in (0, 1)$ and $y \in \mathbb{Q}_+^n$, we must find a configuration in $C^* \in C_J$ with $\sum_{j \in C^*} y_j \geq (1 - \epsilon) \cdot \max_{C \in C_J} (\sum_{j \in C} y_j)$. Recall that the configurations in C_J are exactly the subsets of demands $C \subset [n]$ with $\sum_{j \in C} s_j \leq 1$ and $\sum_{j \in C} s_j + \max_{j \in C} s_j \leq B$. Our algorithm for the subproblem is the following:

1. Guess the index j^* of the largest-sized item used by the configuration achieving the maximum $\max_{C \in C_J} (\sum_{j \in C} y_j)$.
2. Define the knapsack instance with demands of size at most s_{j^*} in $[n] - \{j^*\}$ such that each remaining demand j has size s_j and value y_j . The knapsack size is $\min(1 - s_{j^*}, B - 2s_{j^*})$. Let v^* be the optimal value of this knapsack instance. Run the knapsack FPTAS to obtain a subset of demands $\bar{C} \subset [n] - \{j^*\}$ with value at least $(1 - \epsilon) \cdot v^*$ in time $\text{poly}(n, 1/\epsilon)$.
3. Output the demands $\bar{C} \cup \{j^*\}$.

There are n guesses for j^* , so the algorithm runs in time $\text{poly}(n, 1/\epsilon)$. For correct guess of j^* , the algorithm outputs $\bar{C} \cup \{j^*\}$, which is a feasible configuration by definition of the residual knapsack instance (we only use demands of size at most s_{j^*} , so $\max_{j \in \bar{C} \cup \{j^*\}} s_j = s_{j^*}$ and the knapsack budget ensures the required constraints.) Further, we have $\sum_{j \in \bar{C} \cup \{j^*\}} y_j \geq (1 - \epsilon) \cdot v^* + y_{j^*} \geq (1 - \epsilon) \cdot \max_{C \in C_J} (\sum_{j \in C} y_j)$. Thus, we can solve SUBPROBLEM for LP'_J in time $T(n, \epsilon) = \text{poly}(n, 1/\epsilon)$.

Now we can apply the theorem to LP'_J to obtain an extreme point of LP'_J , say x^* , with $\sum_{c' \in C_J} x^* \leq LP'_J + \delta$ in time $\text{poly}(n, 1/\delta)$. It follows, $2x^*$ is an extreme point of LP_J with objective value at most $2 \cdot LP'_J + \delta = LP_J + \delta$.

Because $2x^*$ is an extreme point of LP_J , which has n non-trivial constraints, $2x^*$ has at most n non-zero variables. Let $\bar{C}_J = \{c' \in C_J \mid 2x^*_{c'} > 0\}$ be the sub-collection of configurations used by this extreme point. Then $|\bar{C}_J| \leq n$. Now, consider modifying LP_J by keeping only the variables indexed by \bar{C}_J . Let the resulting LP be $LP_J(\bar{C}_J)$. We have $LP_J(\bar{C}) \leq LP_J + \delta$ because they share the solution $2x^*$.

Finally, we relate $LP_J(\bar{C}_J)$ with LP_{mach} . First, we map the “per-demand” configurations of \bar{C}_J to the “per-type” configurations of C as follows: Suppose the demands J are partitioned into types $J = \cup_t J_t$. Then For each configuration $c' \in \bar{C}_J$, we map c' to a configuration in C with $|c' \cap J_t|$ -many demands of type t for every type t . Note that this can be done efficiently because we $|\bar{C}_J| \leq n$. By definition, the mapped configuration is feasible. Let $\bar{C} \subset C$ be the all per-type configurations that are mapped to by some configuration in \bar{C}_J . Then $|\bar{C}| \leq n$. Further, we have $LP_{mach}(\bar{C}) = LP_J(\bar{C}_J)$ by an analogous argument as in [Lemma 15](#).

To conclude, let $LP_{mach}(\bar{C})$ be obtained from LP_{mach} by keeping only variables indexed by \bar{C} . Then, $LP_{mach}(\bar{C})$ has polynomially many variables and constraints, so we can explicitly solve $LP_{mach}(\bar{C})$ to obtain an optimal extreme point, which is also an extreme point of LP_{mach} with objective value $LP_{mach}(\bar{C}) = LP_J(\bar{C}_J) \leq LP_J + \delta = LP_{mach} + \delta$, as required. \square