Load balancing under d-thinning [*]

Ohad Noy Feldheim[†]   Jiange Li [‡]

Abstract

In the classical balls-and-bins model, m balls are allocated into n bins one by one uniformly at random. In this note, we consider the d-thinning variant of this model, in which the process is regulated in an on-line fashion as follows. For each ball, after a random bin has been selected, an overseer may decide, based on all previous history, whether to accept this bin or not. However, one of every d consecutive suggested bins must be accepted. The *maximum load* of this setting is the number of balls in the most loaded bin. We show that after $\Theta(n)$ balls have been allocated, the least maximum load achievable with high probability is $(d + o(1)) \sqrt[d]{\frac{d \log n}{\log \log n}}$. This should be compared with the related d-choice setting, in which the optimal maximum load achievable with high probability is $\frac{\log \log n}{\log d} + O(1)$.

1 Introduction

In the classical *balls-and-bins* model, m balls are sequentially allocated into n random bins, each selected uniformly and independently at random. This model has been extensively studied in probability theory, random graph theory, and computer science, and it has found many applications in various areas, such as hashing, load balancing, and resource allocation in parallel and distributed systems (see e.g., [2], [5], [6], [11], [12]).

The *maximum load* is defined as the number of balls in the fullest bin. When $m \geq n \log n$, the maximum load is known to be of $\frac{m}{n} + \Theta\left(\sqrt{\frac{m \log n}{n}}\right)$ with high probability (see, e.g., [10]). In [2], Azar, Broder, Karlin and Upfal studied the *two-choice* variant of the balls-and-bins model, in which for the allocation of each ball, an overseer has the freedom of making a choice between two bins, each selected independently and uniformly at random. They showed that for $m = \Theta(n)$, the greedy strategy of selecting the less loaded bin yields, with high probability, an optimal maximum load of $\frac{\log \log n}{\log 2} + O(1)$, an exponential improvement over the classical balls-and-bins model. Since then, this surprising phenomenon has found many algorithmic applications (see e.g. [7] and references therein).

[†]Hebrew University of Jerusalem, Israel. E-mail: ohad.feldheim@mail.huji.ac.il
[‡]Hebrew University of Jerusalem, Israel. E-mail: jiange.li@mail.huji.ac.il

In [3], with motivations from statistics, Dwivedi, the first author, Gurel-Gurevich and Ramdas considered the following *two-thinning* variant of the balls-and-bins model. In this setting, for each ball, after a uniformly random *primary* allocation has been suggested, an overseer may choose to either accept it, or allocate the ball into a new independent and uniformly random *secondary* allocation. Note that the overseer is oblivious to the secondary allocation before deciding whether to accept the primary allocation or not. Hence, the overseer has less choice in this variant than in the two-choice setting. Equivalently, the two-thinning setting could be described as allowing the overseer to discard allocations on-line as long as no two consecutive allocations are rejected (thus justifying the term two-thinning). This setup arises naturally in a statistical scenario in which one collects samples one-by-one and is allowed to decide whether to keep each sample or not, under the constraint that no two consecutive samples are discarded. In [3], the authors showed that the power of two-thinning could reduce the discrepancy of a sequence of random points independently and uniformly chosen from the interval $[0, 1]$ to be near optimal. The authors also show that a weaker $(1 + \beta)$-thinning setting, where discarding is allowed with probability $\beta < 1$ independently for every sample whose predecessor was not discarded, also obtains near optimal discrepancy.

In [4], the first author and Gurel-Gurevich show that by setting a threshold and accepting the primary allocation if a ball is assigned to a bin in which the number of balls that already accepted their primary allocations is below this threshold, the overseer can achieve, with high probability, an optimal maximum load of $(2 + o(1))\sqrt{\frac{2 \log n}{\log \log n}}$, a polynomial improvement over the one-choice setting. Observe that the two-choice setting with errors (i.e., with constant probability of placing a ball into the unselected bin) and the $(1 + \beta)$-thinning setting provide no asymptotic improvement over the one-choice setting (this follows from [9], see also [3]).

In [2], the d-choice setting, where the overseer is given $d > 2$ choices, was also considered. In this setting, the greedy algorithm yields an optimal maximum load of $\frac{\log \log n}{\log d} + O(1)$ with high probability; that is, compared with the case $d = 2$, the performance only improves by a multiplicative factor for larger values of d.

This paper investigates the analogous *d-thinning* setting for $d > 2$, in which the overseer is allowed to reject up to $d - 1$ consecutive allocations for each ball (see an equivalent description in the abstract). Our result is about the recovery and analysis of the optimality of the ℓ-threshold strategy. This strategy rejects the i-ary allocation of a ball if this ball is assigned to a bin which already contains ℓ balls whose i-ary allocations have been accepted. More details on the ℓ-threshold strategy and its optimality are given in the next section. Our main result is the following.

Theorem 1.1. *Let $\rho > 0$ be a constant. Let f be the $\left(\frac{d \log n}{\log \log n}\right)^{1/d}$-threshold strategy for the d-thinning of $\lfloor \rho n \rfloor$ balls into n bins. Then f is asymptotically optimal, and with high probability yields*

$$\mathrm{MaxLoad}^f(\lfloor \rho n \rfloor) = (d + o(1))\left(\frac{d \log n}{\log \log n}\right)^{1/d}.$$

Clearly, the exponent of the maximum load in the d-thinning setting keeps improving as d grows, but it is never asymptotically better than that in the two-choice setting.

Interestingly, our threshold strategy was used by Adler, Chakrabarti, Mitzenmacher and Rasmussento [1] to obtain a similar tight upper bound (up to a constant factor) in the related model of parallel allocation (which is neither stronger nor weaker than the setting considered here). Nevertheless, we provide here a finer analysis of the threshold strategy and achieve an upper bound which is tight up to the $1 + o(1)$ factor.

The more challenging part is to prove the matching lower bound on $\mathrm{MaxLoad}^f(\lfloor \rho n \rfloor)$, which establishes the optimality of the $\left(\frac{d \log n}{\log \log n}\right)^{1/d}$-threshold strategy. This is achieved

in Section 4, not by directly extending the argument in [4], but rather via a reduction lemma, which reduces a d-thinning problem to a $(d-1)$-thinning problem. After $d-1$ rounds of iteration, the original d-thinning problem is reduced to a one-thinning problem, which is an easy to handle problem in the classical balls-and-bins model.

2 Notation and definitions

Representation of the setting Let $\{Z_1(r)\}_{r \in \mathbb{N}}, \cdots, \{Z_d(r)\}_{r \in \mathbb{N}}$ be d sequences of independent random variables, which are uniformly distributed on the set $[n] := \{1, 2, \cdots, n\}$. Here $\{Z_1(r)\}_{r \in \mathbb{N}}$ is used as a pool of primary allocations, $\{Z_2(r)\}_{r \in \mathbb{N}}$ – as a pool of secondary allocations and, in general, $\{Z_d(r)\}_{r \in \mathbb{N}}$ – as a pool of d-ary allocations. Elements are drawn from these pools one by one as needed, so that if the overseer accepts the primary allocation $Z_1(1)$ of the first ball and rejects the primary allocation $Z_1(2)$ of the second ball, then the bin $Z_2(1)$ is offered as the secondary allocation for the second ball.

A d-thinning strategy A d-*thinning strategy* f is a sequence of Boolean functions $f_{t,i} : [n]^t \to \{0, 1\}$ with $i \in [d-1]$, which takes the final allocations of the first $t-1$ balls and the i-ary allocation of the t-th ball, and returns 0 if this i-ary allocation is accepted and 1 if it is rejected. We also allow $f_{t,i}$ to implicitly depend on any external, potentially random data, as long as this data is independent from all future allocations. For completion we also define $f_{t,d} \equiv 0$, which corresponds to the fact that in our setting, the d-ary allocation cannot be rejected. Let $\{Z(t)\}_{t \in \mathbb{N}}$ denote final allocations assigned by f.

Towards obtaining a formal description of $Z(t)$, set $r_1(t) = t$ and for $i \in \{2, \cdots, d\}$ inductively define

$$r_i(t) = \sum_{k=1}^{t} \prod_{j=1}^{i-1} f_{k,j}(Z(1), \ldots, Z(k-1), Z_j(r_j(k))). \tag{2.1}$$

Thus, $r_i(t)$ represents how many of the first t balls have their first $i-1$ rounds of allocations rejected. For $i \in [d]$ and $r \in \mathbb{N}$, we define

$$t_i(r) = \min\{t \in \mathbb{N} : r_i(t) = r\}, \tag{2.2}$$

which is the unique time at which $Z_i(r)$ is used as an i-ary allocation. For $t \in \mathbb{N}$, the function $\mathrm{Chosen}(t)$ determines which round allocation is accepted for the t-th ball, i.e.,

$$\mathrm{Chosen}(t) = \min\left\{i \in [d] : f_{t,i}(Z(1), \ldots, Z(t-1), Z_i(r_i(t))) = 0\right\}.$$

The *final allocation* assigned by f is now defined via

$$Z(t) = Z_{\mathrm{Chosen}(t)}(r_{\mathrm{Chosen}(t)}(t)).$$

Induced thinning strategy For $j \in [d]$, let f^j be the $(d-j+1)$-thinning strategy induced by f as follows

$$f^j_{r,i} = f_{t_j(r),i+j-1}. \tag{2.3}$$

Particularly, we have $f^1 = f$. The thinning strategy f^j accepts or rejects the i-ary allocation of the r-th ball (with respect to f^j) according to the decision of f on the $(i+j-1)$-ary allocation of the $t_j(r)$-th ball (with respect to f), i.e., the r-th ball whose first $j-1$ allocations we rejected. Observe that the $(d-j+1)$-thinning strategy f^j, relies on information concerning the first $j-1$ rounds of allocations assigned by f. This

is still a valid strategy as these allocations are independent of future allocations from $\{Z_k(r_k(t))\}_{k>j}$ and $\{Z_k(r)\}_{k\in[d],r>r_k(t)}$.

We also introduce the operation

$$f^+ = f^2, \tag{2.4}$$

and observe that for all $j \in [d-1]$ we have $(f^j)^+ = f^{j+1}$.

Load analysis notation For $m \in [n]$ and $t \in \mathbb{N}$, the load of bin m after t balls have been allocated is defined as

$$L_m^f(t) = |\{1 \le k \le t : Z(k) = m\}|. \tag{2.5}$$

Let $L_{i,m}^f(r)$ be the number of balls that accept bin m as their i-ary allocations after $t_i(r)$ balls were allocated, i.e.,

$$L_{i,m}^f(r) = |\{1 \le k \le t_i(r) : \text{Chosen}(k) = i, Z(k) = m\}|. \tag{2.6}$$

Clearly, we have $\sum_{i=1}^d L_{i,m}^f(r_i(t)) = L_m^f(t)$. The maximum load over a subset $S \subseteq [n]$ after t balls have been allocated using the strategy f is defined as

$$\text{MaxLoad}_S^f(t) = \max_{m \in S} L_m^f(t). \tag{2.7}$$

In addition, we define

$$\phi_S^f(t) = \left|\left\{m \in S : L_m^f(t) > 0\right\}\right|, \tag{2.8}$$

which is the number of non-empty bins in S after t balls were allocated, and

$$\psi_S^f(t) = |\{m \in S : \text{there exists } 1 \le k \le t \text{ such that } Z_1(k) = m\}|, \tag{2.9}$$

which counts the number of bins in S offered as primary allocations in the process of allocating of t balls. For simplicity, we drop the subscript S when $S = [n]$.

Asymptotic optimality We say that a d-thinning strategy f is *asymptotically optimal* if for any other d-thinning strategy g, with high probability, as n tends to infinity we have

$$\text{MaxLoad}^f(t) \le (1 + o(1)) \cdot \text{MaxLoad}^g(t).$$

Threshold strategy The ℓ-threshold strategy for $\ell > 0$ is the strategy f that accepts an allocation when the suggested bin contains no more than ℓ balls that were allocated in the i-ary allocations round, i.e.,

$$f_{t,i}\big(Z(1), \ldots, Z(t-1), Z_i(r_i(t))\big) = \mathbb{1}\left(L_{i,Z_i(r_i(t))}^f(r_i(t)) > \ell\right),$$

where $\mathbb{1}(E)$ is the indicator function for the event E.

The following lemmata from [4] and [8] will be of use in our analysis. The first is a comparison bound relating the balls-and-bins model to independent Poisson random variables. We denote by $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Given $x, y \in (\mathbb{N}_0)^n$, we say that $x \le y$ if $x_i \le y_i$ for all $i \in [n]$. A set $S \subset (\mathbb{N}_0)^n$ is called *monotone decreasing (respectively, increasing)* if $x \in S$ implies that $y \in S$ for all $y \le x$ (respectively, $x \le y$).

Lemma 2.1 ([8], Theorem 5.10)**.** *Let $\{X_m\}_{m\in[n]}$ be the number of balls in bins $m \in [n]$ when $\lfloor \theta n \rfloor$ balls are independently and uniformly placed into n bins. Let $\{Y_m\}_{m\in[n]}$ be independent Poisson random variables with parameter θ. For any monotone set $S \subseteq [n]$, we have*

$$\mathbb{P}((X_1, \cdots, X_n) \in S) \le 2\mathbb{P}((Y_1, \cdots, Y_n) \in S).$$

Two additional lemmata from [4] provide concentration bounds for the maximum load over a subset of bins and on the number of bins with load above a certain level in the balls-and-bins process, respectively.

Lemma 2.2 ([4], Lemma 2.2)**.** *Let $\{X_m\}_{m\in[n]}$ be the number of balls in bins $m \in [n]$ when $\lfloor \theta n \rfloor$ balls are independently and uniformly placed into n bins. For $k \in \lfloor \theta n \rfloor$ and $S \subseteq [n]$, we have*

$$\mathbb{P}\left(\max_{m\in S} X_m < k\right) \leq 2\exp\left(-\frac{\theta^k |S|}{ek!}\right).$$

Lemma 2.3 ([4], Lemma 2.3)**.** *Let $\{X_m\}_{m\in[n]}$ be the number of balls in bins $m \in [n]$ when $\lfloor \theta n \rfloor$ balls are independently and uniformly placed into n bins. For any $S \subseteq [n]$, we have*

$$\mathbb{P}\left(|\{m \in S : X_m > 0\}| \leq \frac{\theta|S|}{2e}\right) \leq 2\exp\left(-\frac{\theta^2 |S|}{2e^2}\right).$$

3 MaxLoad upper bound: $\left(\frac{d\log n}{\log\log n}\right)^{1/d}$-threshold strategy

This section is dedicated to establishing a tight upper bound for the maximum load of the ℓ-threshold strategy with $\ell = \left(\frac{d\log n}{\log\log n}\right)^{1/d}$, from which the upper bound in Theorem 1.1 is an immediate consequence. Our proof essentially follows from iterations of the argument in [4]. Observe that with this choice of ℓ, we have

$$\ell^d \log \ell = \left(1 - \frac{\log\log\log n - \log d}{\log\log n}\right)\log n,$$

which implies

$$\ell^{\ell^d} = n^{1-o(1)}. \tag{3.1}$$

Proposition 3.1. *Let $\ell = \left(\frac{d\log n}{\log\log n}\right)^{1/d}$. For any $\epsilon > 0$, there exists $n_0 = n_0(d, \rho, \epsilon)$ such that for all $n > n_0$, the ℓ-threshold strategy f satisfies*

$$\mathbb{P}\left(\mathrm{MaxLoad}^f(\lfloor \rho n \rfloor) > (d+\epsilon)\ell\right) < n^{-\epsilon/3}. \tag{3.2}$$

Proof. By definition, the ℓ-threshold strategy guarantees that $L_{i,m}^f(r) \leq \ell$ for all $i \in [d-1]$, $m \in [n]$ and $r \in \mathbb{N}$. Set $L = (1+\epsilon)\ell$. We have

$$\mathbb{P}\left(\mathrm{MaxLoad}^f(\lfloor \rho n \rfloor) > (d+\epsilon)\ell\right) \leq \mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(r_d(\lfloor \rho n \rfloor)) > L\right). \tag{3.3}$$

We will upper bound this probability conditioned on the event that $r_i := r_i(\lfloor \rho n \rfloor)$ is bounded above by β_i for all $i \in [d]$, where the sequence $\{\beta_i\}_{i\in[d]}$ is chosen along the proof so that this event occurs with high probability. Using the law of total probability, we have

$$\mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(r_d) > L\right) \leq \mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(r_d) > L \mid r_d \leq \beta_d\right) + \mathbb{P}(r_d > \beta_d)$$

$$\leq \mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(\lfloor \beta_d \rfloor) > L\right) + \mathbb{P}(r_d > \beta_d)$$

$$\leq \mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(\lfloor \beta_d \rfloor) > L\right) + \sum_{i=2}^{d} \mathbb{P}(r_i > \beta_i \mid r_{i-1} \leq \beta_{i-1}). \tag{3.4}$$

Next, we define the sequence $\{\beta_i\}_{i\in[d]}$ in an iterative way such that the 2nd term of (3.4) is small. We set $r_1 = \lfloor \rho n \rceil$ and $\beta_1 = \rho n$. To define β_2 and upper bound $\mathbb{P}(r_2 > \beta_2 \mid r_1 \leq \beta_1)$, we introduce independent Poissons $\{Y_k\}_{k\in[n]}$ with parameter $\rho_1 = \beta_1/n$. Define

$$Y = \sum_{k=1}^n \max\{Y_k - \ell, 0\}.$$

By Lemma 2.1, we have

$$\mathbb{P}(r_2 > \beta_2) \leq 2\mathbb{P}(Y > \beta_2).$$

Since

$$\mathbb{E}\left(e^{\max\{Y_1 - \ell, 0\}}\right) \leq 1 + \sum_{k=\ell+1}^{\infty} e^{k-\ell} \cdot e^{-\rho_1} \frac{\rho_1^k}{k!} < 1 + \frac{\rho_1^\ell}{\ell!} < \exp\left(\frac{\rho_1^\ell}{\ell!}\right),$$

we have

$$\mathbb{P}(Y > \beta_2) = \mathbb{P}\left(e^Y > e^{\beta_2}\right) < \exp\left(\frac{n\rho_1^\ell}{\ell!} - \beta_2\right).$$

To make the right-hand side of the above inequality small, we set $\beta_2 = 2n\rho_1^\ell/\ell!$ and obtain

$$\mathbb{P}(r_2 > \beta_2) < 2\exp\left(-\frac{\beta_2}{2}\right).$$

Once β_{i-1} has been defined, we define β_i and upper bound $\mathbb{P}(r_i > \beta_i \mid r_{i-1} \leq \beta_{i-1})$ by repeating the above argument, replacing $\{Y_k\}_{k\in[n]}$ with independent Poisson random variables with parameter $\rho_{i-1} := \beta_{i-1}/n$. Applying Lemma 2.1 to such variables we obtain

$$\mathbb{P}(r_i > \beta_i \mid r_{i-1} \leq \beta_{i-1}) \leq 2\mathbb{P}(Y > \beta_i) < 2\exp\left(\frac{n\rho_{i-1}^\ell}{\ell!} - \beta_i\right).$$

Setting $\beta_i = 2n\rho_{i-1}^\ell/\ell!$, this yields

$$\mathbb{P}(r_i > \beta_i \mid r_{i-1} \leq \beta_{i-1}) < 2\exp\left(-\frac{\beta_i}{2}\right).$$

Putting together our definition of β_i and ρ_i, we obtain the recursive formula

$$\frac{\beta_i}{n} = \frac{2}{\ell!} \cdot \left(\frac{\beta_{i-1}}{n}\right)^\ell$$

with $\beta_1 = \rho n$. By solving this recursion we obtain

$$\beta_d = n\rho^{\ell^{d-1}} \left(\frac{2}{\ell!}\right)^{\sum_{i=0}^{d-2} \ell^i} = n\rho^{\ell^{d-1}} \left(\frac{2}{\ell!}\right)^{\frac{\ell^{d-1}-1}{\ell-1}}. \tag{3.5}$$

Observe that $2/\ell! > \ell^{-(\ell-1)}$ when $\ell \geq 3$ (using the fact that all $k \in \mathbb{N}$ satisfy $k! \leq e\sqrt{k}(k/e)^k$). Thus for $n \geq 3$ we have

$$\beta_d > n\left(\frac{\rho}{\ell}\right)^{\ell^{d-1}} > n\ell^{-2\ell^{d-1}} > n\ell^{-\ell^d} = n^{1-o(1)},$$

where the second and third inequalities hold when $\ell > \max\{2, 1/\rho\}$, and the last identity uses the observation $\ell^{\ell^d} = n^{1-o(1)}$ as stated in (3.1). Since $\{\beta_i\}_{i\in[d]}$ is a decreasing sequence, we deduce the following bound on the second term of (3.4).

$$\mathbb{P}(r_i > \beta_i \mid r_{i-1} \leq \beta_{i-1}) < 2\exp\left(-\frac{\beta_d}{2}\right) < e^{-n^{1-o(1)}}. \tag{3.6}$$

To estimate the first term of (3.4), we consider $\{Y_k\}_{i\in[n]}$, a collection of independent Poisson random variables with parameter $\rho_d = \beta_d/n$ and apply Lemma 2.1 to obtain

$$\mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(\lfloor\beta_d\rfloor) > L\right) \leq 2\mathbb{P}\left(\max_{1\leq k\leq n} Y_k \geq L\right).$$

Notice that

$$\mathbb{P}(Y_1 \geq L) = e^{-\rho_d}\sum_{k=L}^{\infty}\frac{\rho_d^k}{k!} \leq \frac{\rho_d^L}{L!}.$$

Take a union bound to obtain

$$\mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(\lfloor\beta_d\rfloor) > L\right) \leq 2n\frac{\rho_d^L}{L!} = \frac{2n}{L!}\left(\rho^{\ell^{d-1}}\left(\frac{2}{\ell!}\right)^{\frac{\ell^{d-1}-1}{\ell-1}}\right)^L,$$

where the identity follows from our definition $\rho_d = \beta_d/n$ and the formular (3.5). Using the inequality $k! \geq \sqrt{2\pi k}(k/e)^k$ for any $k \in \mathbb{N}$, we have $2/L! < (e/L)^L < 1$ when $\ell > e$ and

$$\left(\frac{2}{\ell!}\right)^{L\cdot\frac{\ell^{d-1}-1}{\ell-1}} < \left(\frac{e}{\ell}\right)^{\ell L\cdot\frac{\ell^{d-1}-1}{\ell-1}} = \left(\frac{e}{\ell}\right)^{(1+\epsilon)\ell^2\cdot\frac{\ell^{d-1}-1}{\ell-1}} < \left(\frac{e}{\ell}\right)^{(1+\epsilon)\ell^d}.$$

Therefore, we end up with

$$\begin{aligned}\mathbb{P}\left(\max_{m\in[n]} L_{d,m}^f(\lfloor\beta_d\rfloor) > L\right) &< n\left(\frac{\rho e}{\ell}\right)^{(1+\epsilon)\ell^d}\\ &< n\ell^{-(1+\epsilon/2)\ell^d}\\ &= \exp\left(\log n - \log n\left(1+\frac{\epsilon}{2}\right)\left(1 - \frac{\log\log\log n - \log d}{\log\log n}\right)\right)\\ &< n^{-\epsilon/3}.\end{aligned} \qquad (3.7)$$

The second inequality holds automatically whenever $\rho e < 1$. Otherwise, it holds for all n large enough to satisfy $\epsilon > \log(\rho e)/(\log\sqrt{\ell} - \log(\rho e)) = \Theta((\log n - \log\log n)^{-1})$. The last inequality holds for all n large enough to satisfy $\epsilon > 7(\log\log\log n - \log d)/\log\log n$. The desired statement (3.2) now follows from (3.3), (3.4), (3.6) and (3.7). $\qquad\square$

Remark 3.2. From the above proof, we see that the parameter ϵ is allowed to depend on n and Proposition 3.1 remains valid as long as $\epsilon > C\log\log\log n/\log\log n$ for an appropriate absolute constant C. Thus, with this choice of ϵ, the upper bound in Theorem 1.1 is an immediate consequence of Proposition 3.1.

4 MaxLoad lower bound: generic thinning strategies

In this section, we prove the following proposition, from which the lower bound in Theorem 1.1 is an immediate consequence.

Proposition 4.1. *For any $\epsilon > 0$, there exists $n_0 = n_0(d, \rho, \epsilon)$ such that for all $n > n_0$, any d-thinning strategy f satisfies*

$$\mathbb{P}\left(\mathrm{MaxLoad}^f(\lfloor\rho n\rfloor) < (d-\epsilon)\ell\right) < \exp\left(-n^{\epsilon/3d}\right). \qquad (4.1)$$

We first prove a reduction lemma which reduces a d-thinning allocation problem to a $(d-1)$-thinning problem. The above statement will follow from iterations of this thinning reduction lemma.

Let us consider the following more general problem of allocating $\lfloor \gamma \rfloor$ balls into n bins using a d-thinning strategy f. Let $s > 0$ and let $S \subseteq [n]$ be a subset. We denote by A^f the event that the maximum load over S is less than s, i.e.,

$$A^f = \left\{ \text{MaxLoad}_S^f(\lfloor \gamma \rfloor) < s \right\}. \tag{4.2}$$

Our goal is to upper bound the probability $\mathbb{P}\left(A^f\right)$. We divide the allocation process into $\lceil s \rceil$ stages, each of which consists of allocating $w = \lfloor \gamma/s \rfloor$ balls except for the final stage in which the remaining balls are allocated. We define the level sets $\left\{ S_k^f \right\}_{k=0,1,\cdots,\lceil s \rceil}$ as

$$S_k^f = \left\{ m \in S : L_m^f(kw) \geq k \right\}, \tag{4.3}$$

where $L_m^f(t)$, defined as per (2.5), represents the load of bin m after t balls have been allocated using a d-thinning strategy f. Given $\{\gamma_k\}_{k=0,1,\cdots,\lceil s \rceil}$, an arbitrary sequence with $\gamma_0 = |S|$, we introduce the following events

$$
\begin{aligned}
E_k^f &= \left\{ \left| S_k^f \right| < \gamma_k \right\}, \\
F_k^f &= \left\{ \text{MaxLoad}_S^f(kw) < s \right\}.
\end{aligned}
\tag{4.4}
$$

Here, E_k^f is the event that after k stages there are less than γ_k bins containing k or more balls, and F_k^f is the event that after k stages the maximum load over S is less than s. Recall that f^+, defined as per (2.4), is the $(d-1)$-thinning strategy induced from f. Our reduction argument utilizes the following events

$$
\begin{aligned}
A_k^{f^+} &= \left\{ \text{MaxLoad}_{S_{k-1}^f}^{f^+}(\lfloor \gamma_k \rfloor) < s - (k-1) \right\}, \\
B_k^f &= \left\{ \psi_{S_{k-1}^f}^f(w) < 2\gamma_k \right\},
\end{aligned}
\tag{4.5}
$$

where $\psi_{S_{k-1}^f}^f(w)$ was defined in (2.9) as the number of bins in S_{k-1}^f selected as primary slots in the allocation of w balls using the d-thinning strategy f. We require the following reduction lemma.

Lemma 4.2. *Let f be a d-thinning strategy. For any $\gamma, s > 0, S \subseteq [n]$ and arbitrary sequence $\{\gamma_k\}_{k=0,1,\cdots,\lceil s \rceil}$ with $\gamma_0 = |S|$, under the notations above, we have*

$$\mathbb{P}\left(A^f\right) \leq \sum_{k=1}^{\lceil s \rceil} \mathbb{P}\left(\left(A_k^{f^+} \cup B_k^f \right) \cap \left(E_{k-1}^f \right)^c \right). \tag{4.6}$$

Proof. For simplicity, we write $A = A^f, S_k = S_k^f, E_k = E_k^f, F_k = F_k^f, A_k = A_k^{f^+}$ and $B_k = B_k^f$. Clearly, $A \subseteq E_{\lceil s \rceil} \cap F_{\lceil s \rceil}$ and $F_k \subseteq F_{k-1}$. From the law of total probability, we obtain

$$\mathbb{P}(E_k \cap F_k) \leq \mathbb{P}(E_{k-1} \cap F_{k-1}) + \mathbb{P}((E_k \cap F_k) \cap E_{k-1}^c).$$

By induction, we thus have,

$$\mathbb{P}(A) \leq \mathbb{P}\left(E_{\lceil s \rceil} \cap F_{\lceil s \rceil}\right) \leq \sum_{k=1}^{\lceil s \rceil} \mathbb{P}(E_k \cap F_k \cap E_{k-1}^c), \tag{4.7}$$

where $E_0 = \emptyset$ as we set $\gamma_0 = |S|$. Notice that the size of S_k is at least the number of bins in S_{k-1} which receive at least one ball in the k-th stage. Hence, we have

$$E_k \subseteq \left\{ \phi_{S_{k-1}}^f(w) < \gamma_k \right\}, \tag{4.8}$$

where $\phi^f_{S_{k-1}}(w)$, defined in (2.8), is the number of non-empty bins in S_{k-1} after allocating w balls using our d-thinning strategy f. Using the inequality

$$\text{MaxLoad}^f_S(kw) \geq \text{MaxLoad}^f_{S_{k-1}}((k-1)w) + \text{MaxLoad}^f_{S_{k-1}}(w)$$

and the fact that $\text{MaxLoad}^f_{S_{k-1}}((k-1)w) \geq k-1$, we have

$$F_k \subseteq \left\{ \text{MaxLoad}^f_{S_{k-1}}(w) < s - (k-1) \right\}. \tag{4.9}$$

From the definitions of r_2 and f^+ in (2.1) and (2.4), respectively, it follows that

$$\begin{aligned}
A^c_k \cap \{r_2(w) > \gamma_k\} &\subseteq \left\{ \text{MaxLoad}^f_{S_{k-1}}(w) \geq s - (k-1) \right\}, \\
B^c_k \cap \{r_2(w) \leq \gamma_k\} &\subseteq \left\{ \phi^f_{S_{k-1}}(w) \geq \gamma_k \right\}.
\end{aligned} \tag{4.10}$$

Putting together (4.8), (4.9) and (4.10), we have

$$E_k \cap F_k \subseteq A_k \cup B_k.$$

This, together with (4.7), proves the reduction lemma. $\qquad\square$

Proof of Proposition 4.1. The proof follows from a multi-round iteration of Lemma 4.2. Recall that f^j with $j \in [d]$, given in (2.3), is the induced $(d-j+1)$-thinning strategy. Particularly, we have $f^1 = f$. We start with $\gamma = \rho n$, $S := [n]$, and $s := s_1 = (d-\epsilon)\ell$. The goal is to upper bound $\mathbb{P}(A)$, where $A := A^{f^1}$ is the event defined as per (4.2). Set $w := \lfloor w_1 \rfloor$ with $w_1 = \gamma/s_1$. Let $k_1 \in \{0, \cdots, \lceil s_1 \rceil\}$. Recall our definitions of $S^{f^1}_{k_1}$ and $E^{f^1}_{k_1}, F^{f^1}_{k_1}$ in (4.3) and (4.4), respectively. For simplicity, we write $S_{1,k_1} = S^{f^1}_{k_1}$, $E_{1,k_1} = E^{f^1}_{k_1}$ and $F_{1,k_1} = F^{f^1}_{k_1}$. The sequence $\{\gamma_{1,k_1}\}_{k_1 \in \{0, \cdots, \lceil s_1 \rceil\}}$ is defined as follows

$$\gamma_{1,k_1} = n \left(\frac{\theta_1}{4e} \right)^{k_1},$$

where $\theta_1 = w_1/n$. We denote $A_{1,k_1} = A^{f^2}_{k_1}$ and $B_{1,k_1} = B^{f^1}_{k_1}$, which are defined as per (4.5), using the fact that $f^+ = f^2$. In these notations, the first subscript (also used as the subscript of k) indicates the first round of allocation (i.e., primary allocation). We apply Lemma 4.2 with aforementioned parameters γ, S, s and $\{\gamma_{1,k_1}\}_{k_1 \in \{0, \cdots, \lceil s_1 \rceil\}}$ to obtain

$$\mathbb{P}(A) \leq \sum_{k_1=1}^{\lceil s_1 \rceil} \mathbb{P}((A_{1,k_1} \cup B_{1,k_1}) \cap E^c_{1,k_1-1}). \tag{4.11}$$

Lemma 2.3 with $\theta = \theta_1$ and the set there $S = S_{1,k_1-1}$ yields

$$\mathbb{P}(B_{1,k_1} \cap E^c_{1,k_1-1}) \leq 2 \exp\left(-\frac{\theta_1^2 |S_{1,k_1-1}|}{2e^2} \right) \leq 2 \exp\left(-\frac{\theta_1^2 \gamma_{1,k_1-1}}{2e^2} \right),$$

where the second inequality uses the fact that, under E^c_{1,k_1-1}, we have $|S_{1,k_1-1}| \geq \gamma_{1,k_1-1}$. This, together with (4.11), yields

$$\mathbb{P}(A) \leq \sum_{k_1=1}^{\lceil s_1 \rceil} \mathbb{P}(A_{1,k_1} \cap E^c_{1,k_1-1}) + 2 \sum_{k_1=1}^{\lceil s_1 \rceil} \exp\left(-\frac{\theta_1^2 \gamma_{1,k_1-1}}{2e^2} \right). \tag{4.12}$$

Our goal is therefore to upper bound $\mathbb{P}(A_{1,k_1} \cap E^c_{1,k_1-1})$. This will be achieved by establishing an upper bound for $\mathbb{P}(A_{1,k_1})$ analogous to (4.12), which involves the $(d-2)$-thinning strategy f^3 instead of the $(d-1)$-thinning strategy f^2. We continue this thinning

reduction procedure for $(d-1)$ rounds until the problem is reduced to the analysis of the 1-thinning strategy f^d, which is a problem of deterministic computation.

Firstly, we inductively define some notations analogous to those we used to obtain (4.12). For $i \in [d]$, we define

$$
\begin{aligned}
s_i &= s_{i-1} - (k_{i-1} - 1), \\
w_i &= \gamma_{i-1,k_{i-1}}/s_i,
\end{aligned}
\tag{4.13}
$$

where $k_{i-1} \in \{0, 1, \cdots, \lceil s_{i-1} \rceil\}$, and the sequence $\{\gamma_{i,k_i}\}_{k_i \in \{0,\cdots,\lceil s_i \rceil\}}$ is defined as follows

$$
\gamma_{i,k_i} = \gamma_{i-1,k_{i-1}} \left(\frac{\theta_i}{4e} \right)^{k_i},
\tag{4.14}
$$

where $\theta_i = w_i/n$. For $i \in [d-1]$, we write $S_{i,k_i} = S_{k_i}^{f^i}$, and $E_{i,k_i} = E_{k_i}^{f^i}, F_{i,k_i} = F_{k_i}^{f^i}$, and $A_{i,k_i} = A_{k_i}^{f^{i+1}}, B_{i,k_i} = B_{k_i}^{f^i}$, which are defined in (4.3), (4.4) and (4.5), respectively, with $f := f^i$ and $S := S_{i-1,k_{i-1}-1}$. Here, A_{i,k_i} is the event that the maximum load over $S_{i,k_i-1} = S_{k_i-1}^{f^i}$ of allocating $\lfloor \gamma_{i,k_i} \rfloor$ balls into n bins using the $(d-i)$-thinning strategy f^{i+1} is below s_{i+1}. To keep consistency, we formally set $s_0 = (d-\epsilon)\ell$, $k_0 = 1$, $\gamma_{0,1} = \rho n$, $S_{0,0} = [n]$ and $A_{0,1} = A$.

We state the general step of the recursion as follows. In the i-th step, our task is to upper bound $\mathbb{P}(A_{i-1,k_{i-1}})$. Analogous to (4.11), we apply Lemma 4.2 with the parameters $\gamma = \lfloor \gamma_{i-1,k_{i-1}} \rfloor, S = S_{i-1,k_{i-1}-1}, s = s_i$ and $\{\gamma_{i,k_i}\}_{k_i \in \{0,\cdots,\lceil s_i \rceil\}}$ to obtain

$$
\mathbb{P}(A_{i-1,k_{i-1}}) \leq \sum_{k_i=1}^{\lceil s_i \rceil} \mathbb{P}((A_{i,k_i} \cup B_{i,k_i}) \cap E_{i,k_i-1}^c).
$$

Lemma 2.3 with $\theta = \theta_i$ and the set there $S = S_{i,k_i-1}$ yields

$$
\mathbb{P}(B_{i,k_i} \cap E_{i,k_i-1}^c) \leq 2\exp\left(-\frac{\theta_i^2 |S_{i,k_i-1}|}{2e^2}\right) \leq 2\exp\left(-\frac{\theta_i^2 \gamma_{i,k_i-1}}{2e^2}\right),
$$

where the second inequality again uses the fact that, under E_{i,k_i-1}^c, we have $|S_{i,k_i-1}| \geq \gamma_{i,k_i-1}$. Analogous to (4.12), we have

$$
\mathbb{P}(A_{i-1,k_{i-1}}) \leq \sum_{k_i=1}^{\lceil s_i \rceil} \mathbb{P}(A_{i,k_i} \cap E_{i,k_i-1}^c) + 2\sum_{k_i=1}^{\lceil s_i \rceil} \exp\left(-\frac{\theta_i^2 \gamma_{i,k_i-1}}{2e^2}\right).
\tag{4.15}
$$

Putting together (4.12) and (4.15), we obtain

$$
\mathbb{P}(A) \leq \sum_{k_1=1}^{\lceil s_1 \rceil} \cdots \sum_{k_{d-1}=1}^{\lceil s_{d-1} \rceil} \mathbb{P}(A_{d-1,k_{d-1}} \cap E_{d-1,k_{d-1}-1}^c) + 2\sum_{i=1}^{d-1} \sum_{k_1=1}^{\lceil s_1 \rceil} \cdots \sum_{k_i=1}^{\lceil s_i \rceil} \exp\left(-\frac{\theta_i^2 \gamma_{i,k_i-1}}{2e^2}\right).
\tag{4.16}
$$

We conclude the proof with the estimate of the right-hand side of (4.16). Firstly, we establish lower bounds for θ_i and γ_{i,k_i}. Recall that $\gamma_{0,1} = \rho n$, $\theta_i = w_i/n$, $w_i = \gamma_{i-1,k_{i-1}}/s_i$ and $s_i = (d-\epsilon)\ell - \sum_{j=0}^{i-1}(k_j - 1)$. For $i \in [d-1]$, we iterate equations (4.13) and (4.14) to obtain

$$
\gamma_{i,k_i} = n\prod_{j=1}^{i} \left(\frac{\theta_j}{4e}\right)^{k_j},
\tag{4.17}
$$

$$
\theta_{i+1} = \frac{1}{s_{i+1}} \prod_{j=1}^{i} \left(\frac{\theta_j}{4e}\right)^{k_j}.
\tag{4.18}
$$

Denote $\tilde{\rho} = \min\{1, \rho\}$. Towards obtaining a lower bound for θ_i, we iterate (4.18) to obtain

$$\frac{\theta_1}{4e} = \frac{\rho}{4e(d-\epsilon)\ell} > \frac{\tilde{\rho}}{4ed\ell},$$

$$\frac{\theta_2}{4e} > \frac{1}{4e(d-\epsilon)\ell} \left(\frac{\rho}{4e(d-\epsilon)\ell}\right)^{k_1} > \left(\frac{\tilde{\rho}}{4ed\ell}\right)^{k_1+1},$$

and repeat this process with the simple fact $s_{i+1} < (d-\epsilon)\ell$ (see (4.13)) to inductively verify that for $i \geq 2$ we have

$$\frac{\theta_{i+1}}{4e} > \left(\frac{1}{4e(d-\epsilon)\ell}\right)^{\prod_{j=2}^{i}(k_j+1)} \left(\frac{\rho}{4e(d-\epsilon)\ell}\right)^{k_1 \prod_{j=2}^{i}(k_j+1)} > \left(\frac{\tilde{\rho}}{4ed\ell}\right)^{\prod_{j=1}^{i}(k_j+1)}.$$

Hence, for any $i \in [d-1]$, we have

$$\frac{\theta_{i+1}}{4e} > \left(\frac{\tilde{\rho}}{4ed\ell}\right)^{\prod_{j=1}^{i}(k_j+1)}, \tag{4.19}$$

$$\gamma_{i,k_i} > n\left(\frac{\tilde{\rho}}{4ed\ell}\right)^{\prod_{j=1}^{i}(k_j+2)}, \tag{4.20}$$

where (4.20) follows from plugging (4.19) into (4.17).

Next, we upper bound each term of the right-hand side of (4.16). We begin with the easier 2nd term.

The 2nd term of (4.16)   Notice that θ_i and γ_{i,k_i} are decreasing in both i and k_i. Using the lower bounds (4.19) and (4.20), we have

$$\sum_{i=1}^{d-1} \sum_{k_1=1}^{\lceil s_1 \rceil} \cdots \sum_{k_i=1}^{\lceil s_i \rceil} \exp\left(-\frac{\theta_i^2 \gamma_{i,k_i-1}}{2e^2}\right) < d(dl)^{d-1} \exp\left(-\frac{\theta_{d-1}^2 \gamma_{d-1,k_{d-1}-1}}{2e^2}\right)$$

$$< d(d\ell)^{d-1} \exp\left(-8n\left(\frac{\tilde{\rho}}{4ed\ell}\right)^{3\prod_{j=1}^{d-1}(k_j+2)}\right)$$

$$< d(d\ell)^{d-1} \exp\left(-8n\left(\frac{\tilde{\rho}}{4ed\ell}\right)^{3\left(\frac{d-\epsilon}{d-1}\ell+3\right)^{d-1}}\right)$$

$$< d(d\ell)^{d-1} \exp\left(-8n\left(\frac{\tilde{\rho}}{4ed\ell}\right)^{3(2\ell+3)^{d-1}}\right)$$

$$= e^{-n^{1-o(1)}}. \tag{4.21}$$

In the 3rd inequality, we use the inequality of arithmetic and geometric means, and the fact that $\sum_{j=1}^{d-1}(k_j+2) \leq (d-\epsilon)\ell + 3(d-1)$, which follows from $s_d = (d-\epsilon)\ell - \sum_{j=0}^{d-1}(k_j-1) > 0$. Observe that the quantity inside the exponential function is of order $n\ell^{-\ell^{d-1}}$ and the fact that $n \geq \ell^{\ell^d}$ as stated in (3.1). Hence, the last identity holds for sufficiently large n depending on d and ρ but not ϵ.

The 1st term of (4.16)   Notice that $A_{d-1,k_{d-1}}$ is the event that the maximum load over $S_{d-1,k_{d-1}-1}$ after throwing $\lfloor \gamma_{d-1,k_{d-1}} \rfloor$ balls into n bins without retry (i.e., 1-thinning) is less than s_d. Apply Lemma 2.2 with $\theta = \gamma_{d-1,k_{d-1}}/n$ and $S = S_{d-1,k_{d-1}-1}$ to obtain

$$\mathbb{P}(A_{d-1,k_{d-1}} \cap E^c_{d-1,k_{d-1}-1}) \leq 2\exp\left(-\left(\frac{\gamma_{d-1,k_{d-1}}}{n}\right)^{s_d} \frac{\gamma_{d-1,k_{d-1}-1}}{es_d!}\right),$$

where we use the fact that, under $E^c_{d-1,k_{d-1}-1}$, we have $|S_{d-1,k_{d-1}-1}| \geq \gamma_{d-1,k_{d-1}-1}$. Using (4.20) and the fact that $\gamma_{d-1,k_{d-1}} < \gamma_{d-1,k_{d-1}-1}$, we have

$$\left(\frac{\gamma_{d-1,k_{d-1}}}{n}\right)^{s_d} \frac{\gamma_{d-1,k_{d-1}-1}}{es_d!} > \frac{n}{es_d!}\left(\frac{\tilde{\rho}}{4ed\ell}\right)^{(s_d+1)\prod_{j=1}^{d-1}(k_j+2)}. \tag{4.22}$$

Using $s_d = (d-\epsilon)\ell - \sum_{j=0}^{d-1}(k_j - 1) > 0$ again, we have $(s_d+1) + \sum_{j=1}^{d-1}(k_j+2) \leq (d-\epsilon)\ell + 3d - 2$. Apply the inequality of arithmetic and geometric means to the exponent of the right-hand side of (4.22) to obtain

$$(s_d+1)\prod_{j=1}^{d-1}(k_j+2) \leq \left(\left(1-\frac{\epsilon}{d}\right)\ell + \frac{3d-2}{d}\right)^d < \left(1 - \frac{2\epsilon}{3d}\right)\ell^d,$$

where the 2nd inequality holds when $\epsilon > 9d/\ell = \Theta\big((\log\log n/\log n)^{-1/d}\big)$. Combining with the fact that $es_d! < (d\ell)^{d\ell}$, we have

$$\left(\frac{\gamma_{d-1,k_{d-1}+1}}{n}\right)^{s_d} \frac{\gamma_{d-1,k_{d-1}}}{es_d!} > \frac{n}{(d\ell)^{d\ell}}\left(\frac{\tilde{\rho}}{4ed\ell}\right)^{-(1-2\epsilon/3d)\ell^d} > n\ell^{-(1-\epsilon/2d)\ell^d},$$

where the 2nd inequality holds when $\epsilon > 7d(\log(4ed) - \log\tilde{\rho})/\log\ell = \Theta(1/\log\log n)$. For such n, we obtain

$$\sum_{k_1=1}^{\lceil s_1 \rceil} \cdots \sum_{k_{d-1}=1}^{\lceil s_{d-1} \rceil} \mathbb{P}(A_{d-1,k_{d-1}} \cap E^c_{d-1,k_{d-1}-1}) \leq (d\ell)^{d-1}\exp\left(-n\ell^{-(1-\epsilon/2d)\ell^d}\right)$$

$$< (d\ell)^{d-1}\exp\left(-n^{\epsilon/2d}\right)$$

$$< \exp\left(-n^{\epsilon/3d}\right). \tag{4.23}$$

In the second inequality, we use the fact that $n \geq \ell^{\ell^d}$ as stated in (3.1). The last inequality holds when $\epsilon > 6d(\log d + \log\log(d\ell))/\log n = \Theta(\log\log\log n/\log n)$.

The proposition follows from (4.16), (4.21) and (4.23). $\qquad\square$

Remark 4.3. Similar to Proposition 3.1, the parameter ϵ in Proposition 4.1 is also allowed to depend on n and Proposition 4.1 is valid as long as $\epsilon > C/\log\log n$ for an appropriate absolute constant C. With this choice of ϵ, the lower bound in Theorem 1.1 is an immediate consequence of Proposition 4.1.

References

[1] Adler, M., Chakrabarti, S., Mitzenmacher, M. and Rasmussen, L.: Parallel randomized load balancing, *Random Struct. Algor.*, **13**(2), (1998), 159–188. MR-1642570

[2] Azar, Y., Broder, A., Karlin, A. and Upfal, E.: Balanced allocations, *SIAM Journal of Computing*, **29**(1), 1999, 180–200. MR-1710347

[3] Dwivedi, R., Feldheim, O. N., Gurel-Gurevich, O. and Ramdas, A.: The power of online thinning in reducing discrepancy, *Probab. Theory Related Fields.*, **174**(1-2), 2019, 103–131. MR-3947321

[4] Feldheim, O. N. and Gurel-Gurevich, O.: The power of thinning in balanced allocation, *Preprint*, arXiv:1807.01132.

[5] Karp, R. M., Luby, M. and Meyer auf der Heide, F.: Efficient PRAM simulation on a distributed memory machine, *Algorithmica*, **16**, 1996, 517–542. MR-1407587

[6] Korst, J.: Random duplicated assignment: An alternative to striping in video servers, In *Proceedings of the 5th ACM International Multimedia Conference*, Seattle, WA, ACM Press, New York, 1997, 219–226.

[7] Mitzenmacher, M.: The power of two choices in randomized load balancing, PhD thesis. University of California, Berkeley, 1996. MR-2695522

[8] Mitzenmacher, M. and Upfal, E.: Probability and computing: Randomized algorithms and probabilistic analysis, Cambridge University Press, 2005. MR-2144605

[9] Peres, Y., Talwar, K. and Wieder, U.: Graphical balanced allocations and the $(1 + \beta)$-choice process, *Random Struct. Algor.*, **47**(4), 2015, 760–775. MR-3418914

[10] Raab, M. and Steger, A.: Balls into bins–a simple and tight analysis, In *Proceedings of the 2nd International Workshop on Randomization and Approximation Techniques in Computer Science*, Barcelona, Spain, 1998. MR-1729169

[11] Sanders, P., Egner, S. and Korst, J.: Fast concurrent access to parallel disks, *Algorithmica*, **35**, 2003, 21–55. MR-1938242

[12] Stemann, V.: Parallel balanced allocations, In *Proceedings of the 8th Annual ACM Symposium on Parallel Algorithms and Architectures*, Padua, Italy, ACM Press, New York, 1996, 261–269.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm