Geometric Generalizations of the Power of Two Choices

John W. Byers* Jeffrey Considine* Michael Mitzenmacher†

ABSTRACT

A well-known paradigm for load balancing in parallel and distributed systems is the "power of two choices," whereby an item is stored at the less loaded of two (or more) random alternative servers. We investigate the power of two choices in natural settings where items and servers reside in a geometric space and each item is associated with the server that is its nearest neighbor. This is the setting for example in the Chord distributed hash table, where the geometric space is determined by clockwise distance on a one-dimensional ring. For example, our analysis shows that when n items are placed at n servers with d choices per item, the maximum load at any server is $\log \log n / \log d + O(1)$ with high probability, only an additive constant more than when servers are chosen uniformly at random. Our proofs are quite general, showing that the power of two choices works under a variety of distributions, with most geometric constructions having at most an additive $O(1)$ penalty. We also show that these techniques still work under highly unbalanced distributions, and give sharp bounds on the necessary number of choices. Finally, we provide simulation results demonstrating the load balance that results as the system size scales into the millions.

Categories and Subject Descriptors

G.2.2 [**Discrete Mathematics**]: Graph Theory—*Network Problems*; G.3 [**Probability and Statistics**]: Probabilistic Algorithms

General Terms

Algorithms, Performance

*Boston University, Dept. of Computer Science. Supported in part by NSF grants ANI-0093296 and ANI-9986397. E-mail: {byers,jconsidi}@cs.bu.edu.

†Harvard University, Div. of Engineering and Applied Sciences. Supported in part by NSF grants CCR-9983832, CCR-0118701, CCR-0121154, and an Alfred P. Sloan Research Fellowship. E-mail: michaelm@eecs.harvard.edu.

Keywords

Load balancing, two choices, distributed hash tables

1. INTRODUCTION

A well-known paradigm for balancing load is the "power of two choices" [1, 11, 12, 20], whereby a ball or item is stored at the less loaded of two (or more) random alternatives, which we refer to variously as bins and servers. These methods are used in standard hashing with chaining to reduce the maximum number of items, or *load*, in a bin with high probability. Two or more hash functions are used to pick candidate bins for each item to be inserted. Prior to insertion, the loads of the candidate bins are compared and the item is inserted into the bin with the least load. To search for an item, the hash functions are applied and each candidate bin is examined to locate the item.

We consider applying this load balancing paradigm in natural settings for parallel and distributed computing where the servers and items are placed in a geometric space, and an item is associated with the server that is its nearest neighbor. After an initial random placement of the servers, each item considers two (or more) random choices for its location; it picks the one where the nearest neighboring server has the least load. In contrast to a typical application of this paradigm, where each of the *servers* is selected uniformly at random, our application picks *locations* in the geometric space uniformly at random. As a result, the probability of an item considering a given server s is non-uniform; rather, it is proportional to the volume of the geometric space that lies closest to s. For example, using Euclidean distance and considering the Voronoi diagram induced by the random placement of a set of servers on the plane, the probability that a randomly selected location is closest to s is proportional to the size of the Voronoi region of s.

Prior to this work, non-uniform distributions of this kind were considered in a preliminary investigation in [3]. There the question asked was whether the power of two choices could handle the logarithmic load imbalances occurring in many distributed hash tables and how choices could be integrated without significantly affecting query performance. We extend this work significantly, both by supplying missing theory and generalizing it to handle a much wider class of problems. Other investigations have considered the use of choices for server (location) placement instead of item placement [8, 15], but we note that these methods only address the maximum expected number of items placed at a server, instead of the maximum actual number of items placed at any server.

Our contributions are threefold. First, we describe motivating applications which lack uniform sampling but could benefit from the power of two choices paradigm. Second, we extend the seminal results of Azar, Broder, Karlin and Upfal [1]. Their result holds in the standard setting where all bins are equally likely to be selected by each hash function; they show in this case that if there are n items, n bins, and $d \geq 2$ hash functions, the maximum load of any bin is only $\log \log n / \log d + O(1)$ with high probability. We show that the same asymptotic result holds for various geometrically derived distributions, with the only change being in the additive $O(1)$ term. This result is somewhat surprising, since there are several bins that are more likely to be chosen at each point by an $O(\log n)$ factor more than the average. We also demonstrate that similar results hold even when the choice distribution is extremely biased, and show sharp thresholds between double logarithmic and polynomial maximum loads. Finally, we present experimental results demonstrating the effectiveness of this approach as the size of the system scales up to millions of items and servers.

1.1 Applications

Our original motivation arises from recent work on distributed hash tables (DHTs), which have been proposed as a fundamental building block for peer-to-peer systems [15, 16, 18, 19, 21]. For example, in the original description of Chord [19], servers are hashed onto a one-dimensional ring; items are also hashed onto the ring, and are associated with their nearest neighbor in the clockwise direction. More generally, using consistent hashing [7], both servers and items are hashed to points in a key space and each item is assigned to the nearest server in this space. The standard construction for DHTs builds upon a consistent hashing approach and adds a routing layer to quickly locate the server responsible for a particular point in the key space. This approach has the advantage that only a small fraction of the items are reassigned when a server joins or leaves. However, this method results in an uneven distribution of items to servers. For many scenarios, including Chord, at least one of n servers will be assigned $\Theta(\log n/n)$ of the key space, and one of the n servers will be assigned $\Theta(1/n^2)$ of the key space with high probability, resulting in large imbalances, and in particular a large maximum load. The standard solution to this imbalance is for each server to simulate $\Theta(\log n)$ virtual servers, thus guaranteeing $\Theta(1/n)$ assignments with high probability [19]. However, this approach is both costly in terms of overlay maintenance and does not avoid the well-known $\Theta(\log n/ \log \log n)$ imbalances when the number of items and servers are similar. The natural application of the power of two choices is to hash each item to two or more points in the key space and pick the less loaded server as described before.

Another application of consistent hashing is the use of geographic hash tables in sensor and other ad-hoc wireless networks [17]. Geographic hash tables hash items to geographic locations and assign them to the sensor physically closest to that location. It follows that each sensor is being assigned a region in the Voronoi diagram determined by all sensor placements. In an idealized scenario where each sensor is placed uniformly at random, the logarithmic imbalances previously noted on the one-dimensional ring are still present.

In both geographic hash tables and distributed hash tables, the actual assignments of the key space may be more unbalanced than predicted by the model. This may simply be due to bad luck in server placement, or, for geographic hash tables in particular, the fact that the uniform placement model is not realistic. Other factors such as adversarial attacks or outages localized in space may further exacerbate the imbalance. With this in mind, we also briefly consider highly unbalanced distributions.

2. THE GENERALIZED LAYERED INDUCTION ARGUMENT

Throughout this work, the *load* of a bin is the number of balls currently assigned to it. Later on, the *size* of a bin will be used to determine its probability, and we will assume that they have been rescaled so that they sum to one. The classic load balancing result for the power of two choices is the following theorem, due to [1].

THEOREM 1. *Let n balls be placed sequentially into n bins as follows. Each ball chooses $d \geq 2$ bins independently and uniformly at random (with replacement), and the ball is placed into the bin with the least load of these d bins at the time of the placement, with ties broken arbitrarily. After all the balls are placed, the maximum load over all n bins is $\log \log n / \log d + O(1)$ with probability $1 - o(1/n)$.*

We consider the following generalization. Let n balls be placed sequentially into n bins as follows. Each ball chooses $d \geq 2$ bins independently (with replacement), so that for all k, the probability of picking one of the k most likely bins is at most $f(k/n)$, and the ball is placed into the bin with the least load of these d bins at the time of the placement, with ties broken arbitrarily. After all the balls are placed, what is the maximum load over all n bins (with probability $1 - o(1/n)$)?

Theorem 1 handled the special case of $f(x) = x$, so each bin was chosen with probability $1/n$. We extend the proof techniques of Azar et al. to handle the generalized case. Interestingly, although in all of our applications the probability of choosing a bin is fixed throughout the process, this is not a requirement for our results.

Let $B(n, p)$ be a Bernoulli random variable with parameters n and p. We make use of the following lemmas.

LEMMA 2. *Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables in an arbitrary domain, and let $Y_1, Y_2, \ldots, Y_n$ be a sequence of binary random variables, with the property that $Y_i = Y_i(X_1, \ldots, X_i)$. If*

$$\mathbf{Pr}(Y_i = 1 \mid X_1, \ldots, X_{i-1}) \leq p,$$

then

$$\mathbf{Pr}\left(\sum_{i=1}^{n} Y_i \geq k\right) \leq \mathbf{Pr}(B(n, p) \geq k).$$

LEMMA 3. *(Chernoff's bound)*

$$\mathbf{Pr}(B(n, p) \geq 2np) \leq \mathrm{e}^{-np/3}.$$

We begin with the following theorem.

THEOREM 4. *Let $f(x) \leq \gamma_1 x \ln(1/x)$ whenever $x \leq 1/\gamma_2$, for constants $\gamma_1 > 0$ and $\gamma_2 > 0$. Let n balls be placed sequentially into n bins as follows. Each ball chooses $d \geq 2$*

bins independently with replacement, so that for all k, the probability of picking one of the k most likely bins is at most $f(k/n)$, and is placed into the bin with the least load of these d bins at the time of the placement, with ties broken arbitrarily. After all the balls are placed, the maximum load over all n bins is $\log\log n/\log d + O(1)$ *with probability* $1 - o(1/n)$.

PROOF. We use the following notation, similar to that of [1]. The state at time t refers to the state of the system immediately after the tth ball is placed. Let the *height* of a ball be one more than the number of balls already in the bin in which the ball is placed. That is, if we think of balls as being stacked in a bin by order of arrival, the height of a ball is its position in that bin's stack. The variable $h(t)$ denotes the height of the tth ball, and $\nu_i(t)$ and $\mu_i(t)$ refer to the number of bins with load at least i and the number of balls with height at least i at time t, respectively. We use ν_i and μ_i for $\nu_i(n)$ and $\mu_i(n)$ when the meaning is clear.

Let

$$\beta_i = \begin{cases} n, & i = 0, \ldots, \lceil\gamma_2\rceil - 1; \\ (1/\gamma_2)n & i = \lceil\gamma_2\rceil; \\ 2n\ f(\beta_{i-1}/n)^d & i > \lceil\gamma_2\rceil. \end{cases}.$$

Intuitively, β_i is meant to be an upper bound of ν_i over the lifetime of the process. Let $\mathcal{E}_i$ be the event that $\nu_i \le \beta_i$. Note that for $i \le \lceil\gamma_2\rceil$, $\mathcal{E}_i$ holds with certainty. We now show that, with high probability, if $\mathcal{E}_i$ holds and $\beta_{i+1} \ge 2\ln n$, then $\mathcal{E}_{i+1}$ holds.

Now fix a value of i. Let Y_t be a binary random variable such that

$$Y_t = 1 \textbf{ iff } h(t) \ge i+1 \textbf{ and } \nu_i(t-1) \le \beta_i.$$

That is, Y_t is 1 if the height of the tth ball is at least $i+1$ and at time $t-1$ there are fewer than β_i bins with load at least i. Notice that for the tth ball to have height at least $i+1$, all d of its choices for bins must have had at least i balls. We can conclude the following. Let ω_j represent the bins selected by the jth ball. Then

$$\mathbf{Pr}(Y_t = 1 \mid \omega_1, \ldots, \omega_{t-1}) \le f(\beta_i/n)^d \stackrel{\text{def}}{=} p_i.$$

Thus, from Lemma 2, we may conclude that

$$\mathbf{Pr}(\textstyle\sum_{i=1}^n Y_t \ge k) \le \mathbf{Pr}(B(n, p_i) \ge k).$$

Conditioned on $\mathcal{E}_i$, we have $\sum Y_t = \mu_{i+1}$. Thus

$$\begin{aligned} \mathbf{Pr}(\nu_{i+1} \ge k \mid \mathcal{E}_i) &\le \mathbf{Pr}(\mu_{i+1} \ge k \mid \mathcal{E}_i) \\ &= \mathbf{Pr}(\textstyle\sum Y_t \ge k \mid \mathcal{E}_i) \\ &\le \frac{\mathbf{Pr}(\sum Y_t \ge k)}{\mathbf{Pr}(\mathcal{E}_i)} \\ &\le \frac{\mathbf{Pr}(B(n, p_i) \ge k)}{\mathbf{Pr}(\mathcal{E}_i)}. \end{aligned}$$

Letting $k = \beta_{i+1}$ in the above, we have that

$$\begin{aligned} \mathbf{Pr}(\nu_{i+1} \ge \beta_{i+1} \mid \mathcal{E}_i) &\le \frac{\mathbf{Pr}(B(n, p_i) \ge 2np_i)}{\mathbf{Pr}(\mathcal{E}_i)} \\ &\le \frac{\mathrm{e}^{-np_i/3}}{\mathbf{Pr}(\mathcal{E}_i)}. \end{aligned}$$

We conclude that

$$\mathbf{Pr}(\neg\mathcal{E}_{i+1} \mid \mathcal{E}_i) \le \frac{1}{n^2\mathbf{Pr}(\mathcal{E}_i)}$$

whenever $p_i \ge 6\ln n/n$. Using the bound

$$\mathbf{Pr}(\neg\mathcal{E}_{i+1}) \le \mathbf{Pr}(\neg\mathcal{E}_{i+1} \mid \mathcal{E}_i)\mathbf{Pr}(\mathcal{E}_i) + \mathbf{Pr}(\neg\mathcal{E}_i),$$

we have

$$\mathbf{Pr}(\neg\mathcal{E}_{i+1}) \le \frac{1}{n^2} + \mathbf{Pr}(\neg\mathcal{E}_i).$$

To finish, let i^* be the smallest value of i for which $p_i < 6\ln n/n$. Following the previous line of reasoning,

$$\begin{aligned} &\mathbf{Pr}(\nu_{i^*+1} \ge 12\ln n \mid \mathcal{E}_{i^*}) \\ &\qquad \le \frac{\mathbf{Pr}(B(n, 6\ln n/n) \ge 12\ln n)}{\mathbf{Pr}(\mathcal{E}_{i^*})} \\ &\qquad \le \frac{1}{n^2\mathbf{Pr}(\mathcal{E}_{i^*})}, \end{aligned}$$

and so

$$\mathbf{Pr}(\nu_{i^*+1} \ge 12\ln n) \le \frac{1}{n^2} + \mathbf{Pr}(\neg\mathcal{E}_{i^*}).$$

Further, since the most likely bin has probability at most $\gamma_1 \ln n/n$,

$$\begin{aligned} &\mathbf{Pr}(\mu_{i^*+2} \ge 2 \mid \nu_{i^*+1} < 12\ln n) \\ &\qquad \le \frac{\mathbf{Pr}(B(n, (12\gamma_1(\ln n)^2/n)^d) \ge 2)}{\mathbf{Pr}(\nu_{i^*+1} < 12\ln n)} \\ &\qquad \le \frac{\binom{n}{2}(12\gamma_1(\ln n)^2/n)^{2d}}{\mathbf{Pr}(\nu_{i^*+1} < 12\ln n)} \end{aligned}$$

by a simple union bound, and thus

$$\begin{aligned} \mathbf{Pr}(\mu_{i^*+2} \ge 2) &\le \binom{n}{2}(12\gamma_1(\ln n)^2/n)^{2d} \\ &\qquad + \mathbf{Pr}(\nu_{i^*+1} \ge 12\ln n). \end{aligned}$$

Combining our bounds, we find that

$$\begin{aligned} \mathbf{Pr}(\mu_{i^*+2} \ge 2) &\le \binom{n}{2}(12\gamma_1(\ln n)^2/n)^{2d} + \frac{i^*+1}{n^2} \\ &= o(1/n). \end{aligned}$$

Thus, $\mathbf{Pr}(\mu_{i^*+2} \ge 2) = o(1/n)$. But if $\mu_{i^*+2} \le 1$, the maximum load is i^*+2. We need that i^* is in fact $\log\log n/\log d + O(1)$. Proving this requires substantially more technical effort than the similar step for the original proof of Azar et al. [1], as the recursion is more complex. The details of the technical manipulation appear in Appendix C. □

The next theorem generalizes Theorem 4 to apply to a variety of random processes.

THEOREM 5. *Suppose that the sizes of n bins are determined by a random process so that the following conditions hold:*

1. *With probability $1 - o(1/n)$, the maximum bin size is at most $\delta_1 \ln n/n$ for some constant $\delta_1 > 0$.*
2. *For suitable constants $\delta_2, \delta_3, \delta_4 > 0$, the number of bins of size at least c/n is bounded by $\delta_2 n\mathrm{e}^{-c/\delta_3}$ with probability $1 - o(1/n^2)$ when c is in the range $2 \le c \le \delta_4 \ln n$.*

Then when n balls are placed sequentially into the n bins, picking each bin with probability proportional to its size and using d choices as in Theorem 1, the maximum load over all n bins is at most $\log\log n/\log d + O(1)$ *with probability* $1 - o(1/n)$.

Proof. By Theorem 4, it suffices to show that the size distribution is bounded with logarithmic imbalance with probability $1 - o(1/n)$. That is, we wish to demonstrate that there exist constants $\gamma_1, \gamma_2 > 0$ such that with probability $1 - o(1/n)$, for all $a \leq n/\gamma_2$, the sum of the largest a bin sizes is at most $\gamma_1 \frac{a}{n} \log \frac{n}{a}$.

Let c_1 be a constant to be specified later with $0 < c_1 < 1$. By the first condition, if $a < n^{c_1}$, then the sum of the a largest sizes is at most $\delta_1 \frac{a}{n} \log n \leq \left(\frac{\delta_1}{1-c_1}\right) \frac{a}{n} \log \frac{n}{a}$ with probability $1 - o(1/n)$. This will serve as a base case for bounding the sizes for the remaining a values.

Now consider the a bins of largest sizee for any a. We first bound the total size of the smallest $a/2$ of these a bins, and then we apply the argument recursively to obtain the final bound. By the second condition, the probability there are more than $\delta_2 n e^{-c/\delta_3}$ bins of size at least c/n is $o(1/n^2)$ when $2 \leq c \leq \delta_4 \ln n$. Rephrasing this statement by letting $b/2 = \delta_2 n e^{-c/\delta_3}$, the probability that there are more than $b/2$ bins of size at least $\delta_3 \ln \frac{2\delta_2 n}{b}/n$ is at most $o(1/n^2)$ for $(2\delta_2) n^{1-\delta_4/\delta_3} \leq b \leq (2\delta_2 e^{-2/\delta_3})n$. Hence, the probability that the sum of the sizes of the smallest $b/2$ of the largest b bins exceeds $(\delta_3/2) * (b/n) * \ln \frac{2\delta_2 n}{b}$ for b in the above range is at most $o(1/n^2)$. Thus, we can bound all but the largest $(2\delta_2) n^{1-\delta_4/\delta_3}$ bins, but these are covered by the base case if we set $c_1 = \delta_4 / 2\delta_3$.

Let j be the smallest integer such that $a/2^j \leq (2\delta_2) n^{1-\delta_4/\delta_3}$ (note $j = O(\log n)$). Let c_2 be the constant $\min(2\delta_2 e^{-2/\delta_3}, 1/e)$. For $a \leq c_2 n$, combining the previous statements describing the recursion and the base case, we have that the sum of the largest a bin sizes is very loosely bounded by

$$\sum_{k=0}^{j} \frac{\delta_3}{2} \frac{a/2^k}{n} \ln\left(\frac{2\delta_2 n}{a/2^k}\right) + \left(\frac{\delta_1}{1-c_1}\right) \frac{a}{n} \ln\left(\frac{n}{a}\right)$$

with probability at least

$$1 - \sum_{k=1}^{j} o(1/n^2) + o(1/n).$$

This probability is $1 - o(1/n)$, and we bound the first component of the sum as follows:

$$\begin{aligned}
&\sum_{k=0}^{j} \left(\frac{\delta_3}{2}\right) \frac{a/2^k}{n} \ln\left(\frac{2\delta_2 n}{a/2^k}\right) \\
&\leq \quad \delta_3 \frac{a}{n} \sum_{k=0}^{\infty} \left(\frac{1}{2^{k+1}} \ln\left(\frac{2\delta_2 n}{a}\right) + \frac{1}{2^{k+1}} \ln 2^k\right) \\
&\leq \quad \delta_3 \frac{a}{n} \left(\ln\left(\frac{2\delta_2 n}{a}\right) + \left(\frac{\ln 2}{2}\right) \sum_{k=0}^{\infty} \frac{k}{2^k}\right) \\
&\leq \quad \delta_3 \frac{a}{n} \ln \frac{n}{a} + O\left(\frac{n}{a}\right) \\
&\leq \quad 2\delta_3 \frac{a}{n} \ln \frac{n}{a}
\end{aligned}$$

Folding this bound back in, we find that for $a \leq c_2 n$, the sum of the a largest bin sizes is at most

$$\left(2\delta_3 + \frac{\delta_1}{1-c_1}\right) \frac{a}{n} \ln\left(\frac{n}{a}\right)$$

with probability $1 - o(1/n)$. □

We can now show that the power of two choices works for the geometric distributions motivated earlier with the following corollaries.

Corollary 6. *Suppose that n points are placed independently and uniformly at random on the boundary of a circle with circumference 1. The n induced arcs correspond to bins. Now balls are placed sequentially into the n bins as follows. Each ball chooses d points independently and uniformly at random from the boundary of the circle; each of the d points corresponds to the bin whose arc contains that point. The ball is placed into the bin with the least load of these d bins at the time of the placement, with ties broken arbitrarily. After all the balls are placed, the maximum load over all n bins is at most* $\log\log n/\log d + O(1)$ *with probability* $1 - o(1/n)$.

Corollary 7. *Suppose that n points are placed independently and uniformly at random in the 2-dimensional unit torus. The n induced Voronoi regions correspond to bins. Now balls are placed sequentially into the n bins as follows. Each ball chooses d points independently and uniformly at random in the unit torus; each of the d points corresponds to the bin whose Voronoi region contains that point. The ball is placed into the bin with the least load of these d bins at the time of the placement, with ties broken arbitrarily. After all the balls are placed, the maximum load over all n bins is at most* $\log\log n/\log d + O(1)$ *with probability* $1 - o(1/n)$.

Proof. (of Corollaries 6 and 7) While the size distributions of each model are already well-known [10, 13], the conditions of Theorem 5 require more careful analysis handling dependencies between bin sizes. This analysis may be found in Appendices A and B, for the random arc model and the random Voronoi cell model, respectively. □

In summary, the logarithmic size imbalances of both geometric models do not affect the load balancing effects of the power of two choices by more than an additive constant. Without choices, both models expect a maximum load of $\Theta(\log n)$ since they have maximum bin sizes of $\Theta(\log n/n)$ with high probability.

2.1 Remarks

We make several remarks regarding these results. First, the constants that arise in the analysis are large for practical considerations, although they could be improved with some further technical work. Simulation results show that the constants are actually quite small in practice. Second, there is a corresponding $\log\log n/\log d - O(1)$ lower bound, which can be proven by following the proof for the standard case [1]. Third, these proofs can be extended to the case where the number of balls is $m \neq n$. Following the argument [1], we find the maximum load is $O(m/n) + O(\log\log n/\log d)$ with high probability. Stronger bounds, following the lines of [2], may be possible. Fourth, Vöcking suggested a variation of the original scheme that achieves an improved upper bound of $\log\log n/d \log \phi_d + O(1)$, where ϕ_d is the asymptotic growth rate of the d-th order Fibonacci numbers, by breaking ties in a novel manner [20]. In this setting, the variation would correspond to each ball picking one point uniformly from each of the d intervals $[0, i/d)$ for $i = 1, 2, \ldots, d$, going to the corresponding least loaded bin, and breaking ties toward the interval with the lowest corresponding value of i. Our proof could be modified to show that the $\log\log n/d \log \phi_d + O(1)$ upper bound holds in this setting with this variation as well. In our experimental section, we describe an alternative tie breaking scheme for non-uniform distributions that appears to provide slightly better performance than even this scheme.

We also note that while Chord [19] was the original motivation for Corollary 6, this corollary also applies to distributed hash tables such as Viceroy [9] and Koorde [6] which respectively use butterfly networks and deBruijn graphs instead of a hypercube. All three use the random arc model to divide the hash space. Theorem 4 also applies to distributed hash tables with more uniform distributions of bin sizes, as in [15, 8], since the function $f()$ is only used as an upper bound on the distribution.

Finally, we point out that the modifications we made to the recurrence of [1] are general. For any function f bounding the bin sizes, a recurrence of the form $\beta_i = nc(f(\beta_{i-1}/n))^d$ can be used. Many probability distributions can be handled in this fashion, and we note that it suffices for $f(x)$ to be a reasonable bound as x approaches zero. The next section gives another example for more highly skewed distributions illustrating the importance of the most likely bins in the analysis.

3. HIGHLY UNBALANCED CASES

As a thought experiment, we now consider what happens when the bin sizes are highly skewed, following a power law distribution. In such a case, the most likely bins may be chosen a polynomial factor more than average. Despite this imbalance, we find that the power of two choices still works if the number of choices is allowed to vary with the power law exponent. Furthermore, for any particular power law exponent, we show a threshold for the number of choices, where above this threshold the maximum load grows double logarithmically with n, and below this threshold, the maximum load grows as a small polynomial in n.

Theorem 8. *Let $f(x) \leq x^\alpha$ where α is a constant and $0 < \alpha < 1$. Let n balls be placed sequentially into the n bins as follows. Each ball chooses $d > 1/\alpha$ bins independently with replacement, in such a fashion that the probability of picking one of the k most likely bins is at most $f(k/n)$, and is placed into the bin with the least load of these d bins at the time of the placement, with ties broken arbitrarily. After all the balls are placed, the maximum load over all n bins is at most $\log\log n/\log(\alpha d) + 1/(\alpha d - 1) + O(1)$ with probability $1 - o(1/n)$ or $\log\log n/\log(\alpha d) + O(1)$ with probability $1 - o(1)$.*

Proof. The proof is similar to those of Theorems 1 and 4 so we only sketch the differences.

First, we use the recurrence

$$B_i = \begin{cases} n, & i = 1, \ldots, 5; \\ n/2e & i = 6; \\ ne(\beta_{i-1}/n)^{\alpha d} & i > 6. \end{cases}$$

which gives

$$\beta_{6+i} = \frac{ne^{((\alpha d)^i - 1)/(\alpha d - 1)}}{(2e)^{(\alpha d)^i}}.$$

This results in $i^* \leq \log\log n/\log(\alpha d) + O(1)$.

As before, we have

$$\mathbf{Pr}(\nu_{i^*+1} \geq 12 \ln n) \leq \frac{1}{n^2} + \mathbf{Pr}(\neg\mathcal{E}_{i^*}).$$

To bound the tail with probability $1 - o(1)$ as in [1], we can use the same approach as for Theorems 1 and 4. To achieve the higher probability bound $1 - o(1/n)$ (at the cost of a looser bound on the maximum load), we use a slightly modified version. Using a union bound,

$$\begin{aligned} \mathbf{Pr}(\mu_{i^*+j+1} \geq 0 \mid \nu_{i^*+j} \leq 12 \ln n) & \\ \leq \quad n \left(\frac{12 \ln n}{n} \right)^{\alpha d} & \end{aligned}$$

and by induction

$$\mathbf{Pr}(\mu_{i^*+j+1} \geq 0 \mid \nu_{i^*} \leq 12 \ln n) \leq \frac{(12 \ln n)^{\alpha d j}}{n^{(\alpha d - 1) j}}.$$

For $j > 1/(\alpha d - 1)$, this is $o(1/n)$. Combining our bounds, we find that

$$\mathbf{Pr}(\mu_{i^*+1/(\alpha d-1)+2} \geq 0) \leq o(1/n),$$

which proves the theorem. ☐

As suggested earlier, a natural followup question to ask is whether Theorem 8 is tight. That is, what happens if $d < 1/\alpha$?

Theorem 9. *If the most likely bin has size $1/n^\alpha$, where $0 < \alpha < 1$ is a constant, and $d < 1/\alpha$, then when the n balls are placed sequentially into the n bins with d choices as in Theorem 1, with probability $1 - o(1/n)$ the number of balls in that bin is at least $\Omega(n^{1-\alpha d})$.*

Proof. By construction, the probability that a particular ball will choose this bin all d times is $n^{-\alpha d}$, and the expected number of balls making these choices is $n^{1-\alpha d}$. Note that $0 < 1 - \alpha d < 1$, so this is a low degree polynomial in n. For large enough n, this is of course more than $2 \log n$ and standard Chernoff bounds apply. ☐

4. EXPERIMENTAL RESULTS

Because the theoretical analysis introduces significant constants, the power of two choices is best tested with experiments. First, we verify the need for better load balancing in practice when $d = 1$. Table 1 shows the maximum observed loads over all trials for varying numbers of bins and choices. For this and later experiments, we ran 1000 trials, and the numbers of balls and bins were the same ($m = n$). Unless otherwise mentioned, all ties are broken randomly. In all cases, there is a large drop in load when the number of choices increases from one to two. All pseudo-random numbers were generated using the well-known linear congruential generator, namely the standard `drand48` and `lrand48` functions, seeded by the current time with millisecond accuracy.

Tables 2 and 3 examine the distribution of the maximum load more closely when there are two or more choices. Each entry lists the relative frequences of the maximum load. For example, in Table 2, the entry for $n = 2^{12}$ and $d = 3$ shows that 88.1% of the trials had a maximum load of 3, while the remaining 11.9% of them had a maximum load of 4. For each choice of parameters, the distribution is tightly concentrated with only two possible outcomes occurring during all but one of our experiments. In that one remaining experiment (using random arcs with 2^{12} balls and bins, and two choices), there were three distinct outcomes. In contrast, the maximum load without choices is highly variable. For example, when repeating the scenario using random arcs with 2^{12} balls and

d	$n = 2^{12}$	$n = 2^{16}$	$n = 2^{20}$	$n = 2^{24}$
2	**4** ... 88.5% **5** ... 11.4% **6** ... 0.1%	**4** ... 20.9% **5** ... 79.1%	**5** ... 100.0%	**5** ... 99.3% **6** ... 0.7%
3	**3** ... 88.1% **4** ... 11.9%	**3** ... 17.4% **4** ... 82.6%	**4** ... 100.0%	**4** ... 100.0%
4	**3** ... 100.0%	**3** ... 99.9% **4** ... 0.1%	**3** ... 99.3% **4** ... 0.7%	**3** ... 86.1% **4** ... 13.9%

Table 2: Experimental maximum load with random arcs ($m = n$)

d	$n = 2^{12}$	$n = 2^{16}$	$n = 2^{20}$	$n = 2^{24}$
2	**3** ... 57.2% **4** ... 42.8%	**4** ... 100.0%	**4** ... 99.6% **5** ... 0.4%	**4** ... 96.0% **5** ... 4.0%
3	**3** ... 100.0%	**3** ... 100.0%	**3** ... 99.3% **4** ... 0.7%	**3** ... 93.2% **4** ... 6.8%
4	**2** ... 30.9% **3** ... 69.1%	**3** ... 100.0%	**3** ... 100.0%	**3** ... 100.0%

Table 3: Experimental maximum load with random Voronoi cells ($m = n$)

d	$n = 2^{12}$	$n = 2^{16}$	$n = 2^{20}$	$n = 2^{24}$
1	21	27	31	32
2	6	5	5	6
3	4	4	4	4
4	3	4	4	4

(a) Random arcs

d	$n = 2^{12}$	$n = 2^{16}$	$n = 2^{20}$	$n = 2^{24}$
1	12	14	17	19
2	4	4	5	5
3	3	3	4	4
4	3	3	3	3

(b) Random Voronoi cells

Table 1: Maximum observed load over 1000 trials. ($m = n$)

bins, but not allowing choices, the maximum load ranged from 8 to 21 over 1000 trials.

As mentioned earlier, breaking ties appropriately can yield better results. Our theoretical results are based on pessimistically assuming that in case of a tie where two or more choices have the same load, the most likely bin (with the largest area) is chosen. A more optimistic approach is to break ties by choosing the least likely bin (with the smallest area), which is less likely to be chosen in the future. Figure 1 compares four natural tie-breaking schemes: *random*, which chooses randomly, *largest* and *smallest*, which choose the bins with the largest and smallest areas respectfully, and *left*, an adaptation of Vöcking's scheme [20]. As expected, *largest* performs the worst, followed by *random*. Somewhat surprisingly, *smallest* performs the best, beating *left*. An interesting open problem would be to determine the exact performance of this variant, or at least to determine whether or not it is provably better asymptotically than breaking ties randomly.

5. CONCLUSION

At the heart of this paper is an interesting generalization of the two-choice paradigm to geometric settings where the probability of choosing a bin is non-uniform. In parallel and distributed computing settings, such scenarios naturally arise when nearest-neighbor methods are used in conjunction with hashing for random placement. This is exactly the setting in a Chord system without the use of virtual servers. We have demonstrated that the benefits of the two-choice paradigm extend to this setting.

While our bounds theoretically show the value of two choices, it would be useful for practical applications if the theory could be used to accurately predict the resulting load distribution. In the case of uniform bin sizes, this can be done quite well using methods based on differential equations pioneered by Mitzenmacher [11]. While not as accurate as differential equations, the witness tree approach, as demonstrated by Vöcking [20], gives a somewhat tighter analysis than the original argument of Azar et al. It is not clear whether either of these methods can be made to apply to this setting.

While we believe the two-choice paradigm will prove useful for Chord-like networks, there is still considerable work to be done determining how to apply it while maintaining reliability and other useful features of these systems [3].

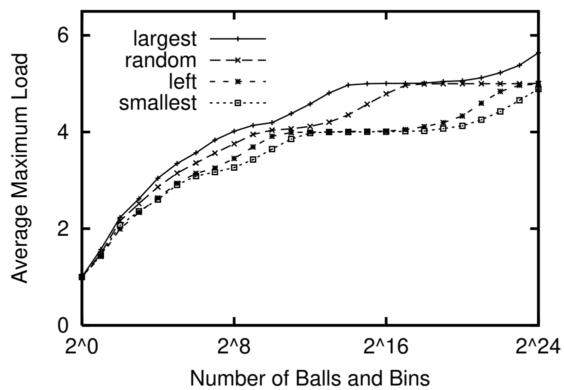
6. REFERENCES

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

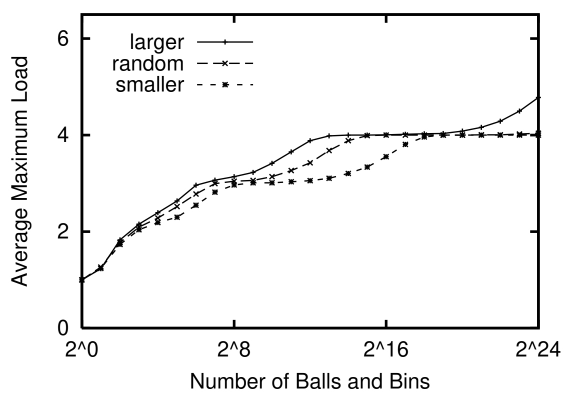
[2] Berenbrink, P., Czumaj, A., Steger, A. and Vöcking, B. Balanced allocations: the heavily loaded case. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, pp. 745-754, 2000.

[3] Byers, J., Considine, J., and Mitzenmacher, M. Simple Load Balancing for Distributed Hash Tables. In *Proceedings of the 2nd International Peer-to-Peer Symposium*, 2003.

[4] Dabek, F., Kaashoek, M. F., Karger, D., Morris, R., and Stoica, I. Wide-area cooperative storage with CFS. In



(a) Random arcs



(b) Random Voronoi cells

Figure 1: Average maximum load varying tie-breaking strategies ($d = 2$, $m = n$)

Proceedings of the 18th ACM Symposium on Operating Systems Principles (SOSP '01) (Chateau Lake Louise, Banff, Canada, Oct. 2001).

[5] Dubhashi, D.P. and Ranjan, D. Balls and bins: a case study in negative dependence. *Random Structures and Algorithms*, 13(2), pp. 99-124, 1998.

[6] Kaashoek, M.F., and Karger, D.R. Koorde: A simple degree-optimal distributed hash table. In *Proceedings of the 2nd International Peer-to-Peer Symposium*, 2003.

[7] Karger, D. R., Lehman, E., Leighton, F. T., Panigrahy, R., Levine, M. S., and Lewin, D. Consistent hashing and random trees: Distributed caching protocols for relieving hot spots on the world wide web. In *ACM Symposium on Theory of Computing* (May 1997), pp. 654–663.

[8] Karger, D.R. and Ruhl, M. Simple Efficient Load Balancing Algorithms for Peer-to-Peer Systems. In *Proceedings of the 3rd International Peer-to-Peer Symposium*, 2004.

[9] Malkhi, D., Naor, M., and Ratajczak, D. Viceroy: A scalable and dynamic emulation of the butterfly. In *Proceedings of PODC '02*, 2002.

[10] Miles, R. E. A synopsis of 'Poisson flats in Euclidean spaces'. In: E. F. Harding and D. G. Kendall (eds), *Stochastic Geometry.* New York, John Wiley, pp. 202-227, 1974.

[11] Mitzenmacher, M. The power of two choices in randomized load balancing. Ph.D. thesis, U.C. Berkeley, 1996.

[12] Mitzenmacher, M., Richa, A., and Sitaraman, R. *The Power of Two Choices: A Survey of Techniques and Results.* Kluwer Academic Publishers, Norwell, MA, 2001, pp. 255–312. edited by P. Pardalos, S. Rajasekaran, J. Reif, and J. Rolim.

[13] Moller, J. Random tesselations in R^d. *Advances in Applied Probability*, 21, pp. 37-73, 1989.

[14] Motwani, R. and Raghavan, P. Randomized Algorithms. Cambridge University Press, 1995.

[15] Naor, M. and Wieder, U. Novel architectures for P2P applications: the continuous-discrete approach. In *Proceedings of SPAA 2003*, pp. 50-60.

[16] Ratnasamy, S., Francis, P., Handley, M., Karp, R., and Shenker, S. A scalable content addressable network. In *ACM SIGCOMM* (2001), pp. 161–172.

[17] Ratnasamy, S., Karp, B., Shenker, S., Estrin, D., Govindan, R., Yin, L., and Yu, F. Data-Centric Storage in Sensornets with GHT, A Geographic Hash Table. In *ACM MONET*, (2003).

[18] Rowstron, A., and Druschel, P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In *Proceedings of Middleware 2001* (2001).

[19] Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., and Balakrishnan, H. Chord: A scalable peer-to-peer lookup service for internet applications. In *ACM SIGCOMM* (2001), pp. 149–160.

[20] Vöcking, B. How asymmetry helps load balancing. In *Proceedings of the* 40^{th} *IEEE-FOCS* (1999), pp. 131–140.

[21] Zhao, B. Y., Huang, L., Stribling, J., Rhea, S. C., Joseph, A. D., and Kubiatowicz, J. D. Tapestry: A Resilient Global-scale Overlay for Service Deployment. *IEEE Journal on Selected Areas in Communications 22*, 1 (2004), 41–53.

APPENDIX

A. CONDITIONING FOR RANDOM ARCS

We now prove Corollary 6 by demonstrating that the random arc model of Chord satisfies the necessary conditions of Theorem 5. We use the Chernoff bound to bound the number of arcs of length at least x for various values of x. Unfortunately, the arc lengths are *dependent* random variables, so the Chernoff bound cannot immediately be applied. One way to cope with this dependence is to use a martingale argument. While the obvious martingale argument gives a weaker bound, it nonetheless suffices for our main result. However, we have found that in the case of arc lengths we can also make use of *negative dependence* [5].

For the purposes of this paper, we say that a collection of 0-1 random variables $X_1, X_2, \ldots, X_n$ is *negatively dependent* if and only if

$$\mathbf{E}\left[X_{i_1} X_{i_2} \cdots X_{i_k}\right] \leq \mathbf{E}\left[X_{i_1}\right] \mathbf{E}\left[X_{i_2}\right] \cdots \mathbf{E}\left[X_{i_k}\right]$$

for any distinct indices $i_1, i_2, \ldots, i_k$ in the range $[1, n]$. Let $X = \sum_{i=1}^{n} X_i$. A simple consequence of the above fact, since each variable takes only the values 0 or 1, is

$$\mathbf{E}\left[\mathrm{e}^{tX}\right] \leq \prod_{i=1}^{n} \mathbf{E}\left[\mathrm{e}^{tX_i}\right].$$

(This can be seen by expanding e^{tX} using the Taylor series expansion, and using $\mathbf{E}\left[X_{i_1}^j\right] = \mathbf{E}\left[X_{i_1}\right]$ since all variables are 0-1.) This is all that is required for the Chernoff bound for the upper tail to hold, following the standard proof [14]. Hence, we can apply the Chernoff bound of Lemma 3 as long as our random variables are negatively dependent.

We now prove some key lemmata regarding the distribution of the arc lengths when n points are thrown randomly onto the circle, as well as negative dependence between sufficiently long arc lengths.

LEMMA 10. *Suppose n points are thrown independently and uniformly at random on the boundary of the unit circle. Let $Z_j = 1$ if the counterclockwise arc from the jth point has length at least c/n, and $Z_j = 0$ otherwise. Then the Z_j are negatively dependent.*

PROOF. Without loss of generality, consider the k random variables $Z_1, Z_2, \ldots, Z_k$. We have $\mathbf{E}[Z_i] = \left(1 - \frac{c}{n}\right)^{n-1}$ for all i, and hence

$$\prod_{i=1}^{k} \mathbf{E}[Z_i] = \left(1 - \frac{c}{n}\right)^{k(n-1)}.$$

Now $Z_1 Z_2 \ldots Z_k = 1$ if and only if all of the arcs associated with each of the k points has length at least c/n; we bound the probability of this happening. Consider the points being placed one at a time. We first consider the first k points, which require some care.

For $1 \le i \le k$, let A_i be the event that the first i points each have non-overlapping arcs of length at least c/n. Let B_i be the event that the ith point does not fall within c/n (in counterclockwise distance) of any of the first $(i-1)$ points. Let C_i be the event that the arc of length c/n from the ith point does not contain any of the first $(i-1)$ points. Clearly, $\mathbf{Pr}(A_1) = 1$, and

$$\begin{aligned}
&\mathbf{Pr}(A_{i+1}) \\
&\quad = \mathbf{Pr}(A_i) \cdot \mathbf{Pr}(B_{i+1} \mid A_i) \cdot \mathbf{Pr}(C_{i+1} \mid A_i B_{i+1}).
\end{aligned}$$

The probability $\mathbf{Pr}(B_{i+1} \mid A_i)$ is easy to calculate as the arcs are necessarily disjoint.

$$\mathbf{Pr}(B_{i+1} \mid A_i) = 1 - \frac{ci}{n} \le \left(1 - \frac{c}{n}\right)^{i}.$$

To compute $\mathbf{Pr}(C_{i+1} \mid A_i B_{i+1})$, we examine an equivalent view of the experiment. Consider starting with a circle of circumference $1 - i(c/n)$ and placing the $i+1$ points uniformly at random. Then think of expanding each of the first i points into an empty arc of length c/n in the counterclockwise direction, giving a corresponding circle of length 1 satisfying events A_i and B_{i+1}. The distribution of point placements we obtain in this fashion is equivalent to that of the original placement approach when conditioned on events A_i and B_{i+1}. The advantage of this point of view is that if we think of the $(i+1)$st point as being placed first, then it is clear that

$$\begin{aligned}
\mathbf{Pr}(C_{i+1} \mid A_i B_{i+1}) &= \left(\frac{1 - c(i+1)/n}{1 - ci/n}\right)^{i} \\
&\le \left(1 - \frac{c}{n}\right)^{i}.
\end{aligned}$$

Combining the above, we find that the probability that after the placement of the first k points there are k disjoint arcs of length at least c/n is at most

$$\begin{aligned}
&\prod_{i=1}^{k-1} \mathbf{Pr}(B_{i+1} \mid A_i) \cdot \mathbf{Pr}(C_{i+1} \mid A_i B_{i+1}) \\
&\qquad\qquad \le \prod_{i=1}^{k-1} \left(1 - \frac{c}{n}\right)^{2i} \\
&\qquad\qquad = \left(1 - \frac{c}{n}\right)^{k^2 - k}.
\end{aligned}$$

For the remaining $n-k$ points, the probability that they miss the k disjoint arcs of length c/n corresponding to the first k points is

$$\left(1 - \frac{ck}{n}\right)^{n-k} \le \left(1 - \frac{c}{n}\right)^{k(n-k)}.$$

Hence the probability that $Z_1 Z_2 \ldots Z_k = 1$, or $\mathbf{E}[Z_1 Z_2 \ldots Z_k]$, is bounded above by

$$\left(1 - \frac{c}{n}\right)^{k(n-k)} \left(1 - \frac{c}{n}\right)^{(k^2-k)} = \left(1 - \frac{c}{n}\right)^{k(n-1)},$$

proving the negative dependence. □

LEMMA 11. *Let N_c be the number of arcs of length at least c/n. For $n \ge c \ge 2$,*

$$\mathbf{Pr}(N_c \ge 2n\mathrm{e}^{-c}) \le \mathrm{e}^{-n\mathrm{e}^{-c}/3}.$$

PROOF. Again let $Z_j = 1$ if the jth point is the starting endpoint of an arc of length at least c/n, and 0 otherwise. From the result in Lemma 10, there is negative dependence between the Z_j's; therefore, Chernoff bounds apply. The probability $Z_j = 1$ can be bounded above by

$$\left(1 - \frac{c}{n}\right)^{n-1} \le \mathrm{e}^{-c}$$

where the inequality holds for $c \ge 2$. Hence $\mathbf{E}[N_c] = \sum_{j=1}^{n} Z_j \le n\mathrm{e}^{-c}$. Applying the standard Chernoff bound of Lemma 3 yields $\mathbf{Pr}(N_c \ge 2n\mathrm{e}^{-c}) \le \mathrm{e}^{-n\mathrm{e}^{-c}/3}$. □

B. CONDITIONING FOR RANDOM VORONOI REGIONS

We now prove Corollary 7 by demonstrating that the random Voronoi region model satisfies the necessary conditions of Theorem 5. It is easy to show the maximum size cell in a Voronoi diagram is bounded by $c_1 \ln n$ for some constant c_1 with high probability. We therefore focus on proving a lemma of the same form as Lemma 11 for this setting. We consider the case of two dimensions with wraparound (i.e. a torus), although it is clear that our argument generalizes to obtain similar bounds with different constant terms for higher constant dimension. Similarly, our results can also be made to apply to the unit circle or unit square; we use the torus to avoid technicalities introduced by boundaries. There has been a fair amount of work on the distribution of the area of regions in random Voronoi diagrams, particularly in two dimensions; see, for example, the work by Miles [10] or Moller [13]. Knowing the distribution is insufficient for our purposes however, since the areas of the Voronoi regions are potentially dependent in non-trivial ways. Interestingly, we have not found tail bounds of the sort we need in previous literature.

By making use of an appropriate set of random variables, we can develop simple bounds via martingales that prove effective for this setting. We begin with a key geometric lemma. Consider any point u in the torus. Let the circular area of area c/n around u be divided into six subregions, as in Figure 2, each of size $c/6n$. Specifically, taking 0 degrees to be parallel to the x-axis and to the right of u, the first subregion is the area from 0 to 60 degrees, and so on.

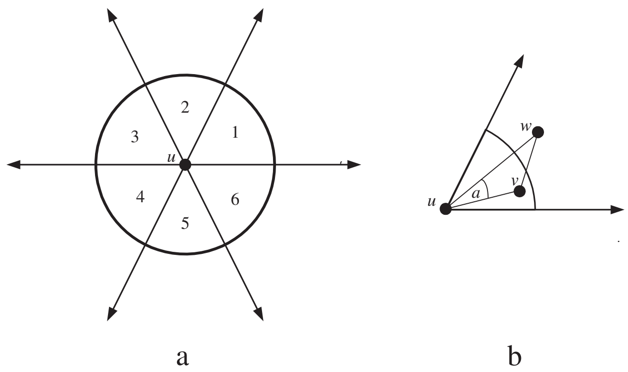


Figure 2: (a) 6 subregions of a circular region around a point u, each corresponding to a 60 degree arc. (b) If v lies in subregion 1 of u, then v is closer to w than u.

Lemma 12. *In a Voronoi diagram derived from n points, if the area of a Voronoi region for a point u is at least c/n, then at least one of the six subregions around u described above does not contain another of the $n-1$ points.*

Proof. Suppose without loss of generality the first subregion of area $c/6n$ around a point u has another of the other $n-1$ randomly placed points, call it v, lying within it. Then we claim that any point in the torus making an angle of between 0 and 60 degrees with u and the x-axis outside of this subregion (and within some fixed radius, say 1/4) is closer to v than to u, and hence not in the Voronoi region of u. This follows clearly from the diagram in Figure 2. Letting w be such a point in the torus, then as in Figure 2, we have

$$d^2_{v,w} = d^2_{u,w} + d^2_{u,v} - 2d_{u,w}d_{u,v}\cos a.$$

The angle a must be between 0 and 60 degrees and $\cos a$ is therefore greater than 1/2. Further, $d_{u,w} > d_{u,v}$. Hence $d^2_{v,w} < d^2_{u,w}$.

We can conclude that if all six of the subregions around u contain another of the $n-1$ points, then the Voronoi region of u is contained within the circle of area c/n around u. This yields the lemma. $\square$

Using this lemma, we show the following.

Lemma 13. *When n points are placed independently and uniformly at random in the unit torus, the number of Voronoi regions of size at least c/n when $\ln n \geq c \geq 12$ exceeds $12ne^{-c/6}$ with probability $o(1/n^4)$.*

Proof. Let $Z_{i,j}$, $1 \leq i \leq n$ and $1 \leq j \leq 6$, be 1 if the jth subregion of area $c/6n$ around the ith point does not contain any of the other $n-1$ points and 0 otherwise. By Lemma 12, $Z = \sum_{i,j} Z_{i,j}$ is clearly an upper bound on the number of Voronoi regions of size at least c/n. Now

$$\mathbf{E}[Z] = 6n\left(1 - \frac{c}{6n}\right)^{n-1} < 6n\mathrm{e}^{-c/6}.$$

Consider the n random points being placed in the circle one at a time. Let X_i be the location of the ith point, and $Y_i = \mathbf{E}[Z \mid X_1, X_2, \ldots, X_i]$. Then the sequence of Y_i is a Doob martingale with expectation $\mathbf{E}[Z]$. Unfortunately, it is not a useful Doob martingale in the context of using Azuma's inequality, since it does not obey a Lipschitz condition. That is, the introduction of a single point X_k can affect a significant number of the random variables $Z_{i,j}$.

To account for this difficulty, we introduce the following modification. Let $F = f(X_1, X_2, \ldots, X_n)$ be the total number of *empty-or-rare* subregions of area $c/6n$ around one of the X_i satisfying the following: either the subregion is empty, or for every X_j in the subregion, there exists $\ln^3 n$ points X_k with $k < i$ and X_j is in the circle of area c/n around X_k. Or, thinking of it another way, as we lay the points in order, each point can change the empty-or-rare status of at most $\ln^3 n$ subregions of other points, and they affect the subregions of the first $\ln^3 n$ points in the order of placement. The idea behind this change is that the rare regions are sufficiently rare that in almost all cases F and Z will be equal, since with high probability no point will ever affect the subregions of $\ln^3 n$ other points. In fact, as $F \geq Z$, we have

$$\mathbf{Pr}(Z \geq 12n\mathrm{e}^{-c/6}) \leq \mathbf{Pr}(F \geq 12n\mathrm{e}^{-c/6}),$$

so it suffices to bound the latter probability.

If we now let $Y_i = \mathbf{E}[F \mid X_1, X_2, \ldots, X_i]$, then the sequence of Y_i is a Doob martingale with expectation $\mathbf{E}[F]$. Now we can apply Azuma's inequality, since the placement of any point can affect only at most $\ln^3 n + 6$ subregions ($\ln^3 n$ subregions of other points, and the six subregions of the point itself). Hence this martingale does obey a Lipschitz condition with bound $\ln^3 n + 6$.

It remains to bound $\mathbf{E}[F]$. Clearly

$$\mathbf{E}[F] \leq \mathbf{E}[Z] + 6n\mathbf{Pr}(Z \neq F),$$

since when F and Z are not the same, the maximum possible value for F is 6n. To bound $\mathbf{Pr}(Z \neq F)$, note that for Z and F to differ, some point X_k must affect the value of some $Z_{i,j}$ for at least $\ln^3 n$ neighboring points. But this means there must be at least $\ln^3 n$ points in an area of size c/n around X_k, where $c \leq \ln n$. Using the standard Chernoff bound from [14], we have that the probability that this happens for any specific X_k is $n^{-O(\log\log n)}$, and hence $\mathbf{Pr}(Z \neq F) = n^{-O(\log\log n)}$. It follows easily that $\mathbf{E}[F] \leq 6n\mathrm{e}^{-c/6}$ as well for n sufficiently large. (That is, $\mathbf{E}[Z] = 6n\left(1 - \frac{c}{6n}\right)^{n-1}$, which differs from $6n\mathrm{e}^{-c/6}$ by more than $n^{-O(\log\log n)}$.) Azuma's inequality [14] yields

$$\mathbf{Pr}(F \geq 12n\mathrm{e}^{-c/6}) \leq \mathrm{e}^{-18n\mathrm{e}^{-c/3}/(\ln^3 n+6)}.$$

For c in the given range, this probability is certainly $o(1/n^4)$. $\square$

C. BOUND FOR THEOREM 4

We now prove the following technical claim, used in the proof of Theorem 4.

Claim 14. *We have $i^* = \frac{\log\log n}{\log d} + O(1)$.*

PROOF. Without loss of generality to Theorem 4, we assume $\gamma_1 \geq 2$ (this loosens the bounds of f). To prove the claim, first note that

$$\beta_{i+1} \geq \frac{\beta_i^d}{n^{d-1}},$$

so a simple induction yields that

$$\ln \frac{n}{\beta_{i+j}} \leq d^i \ln \frac{n}{\beta_j}.$$

We now claim that, using the above, a simple induction shows that for $k \geq 2$

$$\begin{aligned}&\beta_{k+\lceil\gamma_2\rceil} \\ \leq \quad & \left[n \left(\frac{\beta_{\lceil\gamma_2\rceil}}{n}\right)^{d^k} \gamma_1^{d^k+1+2\sum_{j=1}^{k-1} d^j} \right] * \\ & \left[\left(\ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^{\sum_{j=1}^{k} d^j} d^{\sum_{j=1}^{k-1}(k-j)d^j} \right].\end{aligned} \tag{1}$$

To see this, note that

$$\begin{aligned}\beta_{1+\lceil\gamma_2\rceil} &= 2n\left(\gamma_1 \frac{\beta_{\lceil\gamma_2\rceil}}{n} \ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^d \\ &\leq n\left(\frac{\beta_{\lceil\gamma_2\rceil}}{n}\right)^d \gamma_1^{d+1} \left(\ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^d,\end{aligned}$$

and

$$\begin{aligned}\beta_{2+\lceil\gamma_2\rceil} &= 2n\left(\gamma_1 \frac{\beta_{1+\lceil\gamma_2\rceil}}{n} \ln \frac{n}{\beta_{1+\lceil\gamma_2\rceil}}\right)^d \\ &\leq 2n\left(\gamma_1 \frac{\beta_{1+\lceil\gamma_2\rceil}}{n} d \ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^d \\ &= n\left(\frac{\beta_{\lceil\gamma_2\rceil}}{n}\right)^{d^2} \gamma_1^{d^2+2d+1} \left(\ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^{d^2+d} d^d,\end{aligned}$$

giving the base case. The induction is now clear using

$$\begin{aligned}\beta_{k+\lceil\gamma_2\rceil} &= 2n\left(\gamma_1 \frac{\beta_{(k-1)+\lceil\gamma_2\rceil}}{n} \ln \frac{n}{\beta_{(k-1)+\lceil\gamma_2\rceil}}\right)^d \\ &\leq 2n\left(\gamma_1 \frac{\beta_{(k-1)+\lceil\gamma_2\rceil}}{n} d^{k-1} \ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^d\end{aligned}$$

and plugging in equation (1). Bounding the terms of (1) with more pleasant expressions, we have

$$\beta_{k+\lceil\gamma_2\rceil} \leq n\left(\frac{\beta_{\lceil\gamma_2\rceil}}{n}\right)^{d^k} \gamma_1^{3d^k} \left(\ln \frac{n}{\beta_{\lceil\gamma_2\rceil}}\right)^{2d^k} d^{4d^{k-1}}.$$

It follows since $\beta_{\lceil\gamma_2\rceil} = n/\lceil\gamma_2\rceil$ that

$$\beta_{k+\lceil\gamma_2\rceil} \leq n\left(\frac{\gamma_1^3 d^{4/d} \ln^2\lceil\gamma_2\rceil}{\lceil\gamma_2\rceil}\right)^{d^k}.$$

We now choose an appropriate lower bound on γ_2 (using only γ_1) so that

$$\beta_{k+\lceil\gamma_2\rceil} \leq nc^{d^k}$$

for some constant $c < 1$. That i^* is $\frac{\log\log n}{\log d} + O(1)$ now follows. $\square$