

Probabilistic Recurrence Relations for Work and Span of Parallel Algorithms

Joseph Tassarotti

Computer Science Department, Carnegie Mellon University

`jtassaro@cs.cmu.edu`

Abstract

In this paper we present a method for obtaining tail-bounds for random variables satisfying certain probabilistic recurrences that arise in the analysis of randomized parallel divide and conquer algorithms. In such algorithms, some computation is initially done to process an input x , which is then randomly split into subproblems $h_1(x), \dots, h_n(x)$, and the algorithm proceeds recursively in parallel on each subproblem. The total *work* on input x , $W(x)$, then satisfies a probabilistic recurrence of the form $W(x) = a(x) + \sum_{i=1}^n W(h_i(x))$, and the *span* (the longest chain of sequential dependencies), satisfies $S(x) = b(x) + \max_{i=1}^n S(h_i(x))$, where $a(x)$ and $b(x)$ are the work and span to split x and combine the results of the recursive calls.

Karp has previously presented methods for obtaining tail-bounds in the case when $n = 1$, and under certain stronger assumptions for the work-recurrence when $n \geq 2$, but left open the question of the span-recurrence. We first show how to extend his technique to handle the span-recurrence. We then show that in some cases, the work-recurrence can be bounded under simpler assumptions than Karp's by transforming it into a related span-recurrence and applying our first result. We demonstrate our results by deriving tail bounds for the work and span of quicksort and the height of a randomly generated binary search tree.

1 Introduction

The analysis of sequential divide-and-conquer algorithms involves analyzing recurrences of the form

$$W(x) = a(x) + \sum_{i=1}^n W(h_i(x)) \tag{1}$$

where $a(x)$ is the “toll” cost to initially process an input of size x and split it into smaller inputs $h_1(x), \dots, h_n(x)$, which are then processed recursively. When the functions h_1, \dots, h_n are deterministic, there are a number of well-known “cook-book” methods that can be used to obtain asymptotic bounds in many cases [1, 6, 14].

In algorithms where the recursive calls are done in parallel, another measure of cost becomes important: the *span* [3, 4] measures the longest sequential dependency in the execution, and is important for measuring the speed-up of the algorithm as the number of processors is increased. The recurrence for the span has the form:

$$S(x) = b(x) + \max_{i=1}^n S(h_i(x)) \tag{2}$$

where $b(x)$ is the span for processing and dividing the input, but we take the maximum of the recursive calls instead of summing because we are interested in the longest dependency. These kinds of recurrences involving maxima also arise when analyzing heights of search trees or stack space usage of sequential algorithms. Again, when the h_i are deterministic, the above recurrence often simplifies to something of the form $S(x) = b(x) + S(h_i(x))$ for some h_i , and so the usual cook-book methods can be used to obtain asymptotic bounds.

However, when the $h_i(x)$ are random variables (either because the algorithm uses randomness when dividing the input into subproblems, or when considering average case complexity), then $W(x)$ and $S(x)$ are themselves random variables and the situation is more complicated. First, when considering the expected value of the work, we can at least use linearity of expectation to get

$$E[W(x)] = a(x) + \sum_{i=1}^n E[W(h_i(x))] \tag{3}$$

and a variety of techniques can be used to solve the resulting recurrence. But in the case of span, expectation does not commute with taking a maximum, so that in general $E[\max_{i=1}^n S(h_i(x))] \neq \max_{i=1}^n E[S(h_i(x))]$, and we merely have $E[\max_{i=1}^n S(h_i(x))] \geq \max_{i=1}^n E[S(h_i(x))]$. As a result, the expected value of the span does not obey a simple recurrence. A second source of complexity is that we may be interested not just in obtaining bounds on expectation, but may also wish to know that with high probability the work and span are close to this expected value.

Although a number of methods are known for obtaining results for certain recurrences of these forms [7, 8, 9, 11], these often rely on sophisticated tools from probability theory or analytic combinatorics and are not simple cook-book results that can be easily applied.

In contrast to some of the more advanced techniques mentioned above, Karp [12] has presented relatively simple methods for obtaining tail-bounds for work recurrences of the form in Equation 1:

Theorem 1 ([12, Theorem 1.2]). *Suppose $n = 1$ in Equation 1, so that $W(x) = a(x) + W(h(x))$ where $h(x)$ is a random variable such that $0 \leq h(x) \leq x$. Furthermore, assume $E[h(x)] \leq m(x)$ where $0 \leq m(x) \leq x$ and $m(x)/x$ is non-decreasing. Suppose u is a solution to the recurrence $u(x) = a(x) + u(m(x))$. If a and u are continuous and a is strictly increasing on the set $\{x \mid a(x) > 0\}$, then for all positive integers w ,*

$$\text{Prob}[W(x) > u(x) + wa(x)] \leq \left(\frac{m(x)}{x}\right)^w$$

In other words, if there is a single recursive call and we can suitably bound the expected size of $h(x)$, this lets us derive tail-bounds for the work in terms of a simple deterministic recurrence. Often $m(x)/x$ will be a constant fraction such as $1/2$ or $3/4$, so that the recurrence for u can be solved straightforwardly. A second result covers the case where there is more than one recursive call:

Theorem 2 ([12, Theorem 1.5]). *Let W obey the recurrence in Equation 1. Suppose that for all (y_1, \dots, y_n) in the support of the joint distribution $(h_1(x), \dots, h_n(x))$, $E[W(x)] \geq \sum_{i=1}^n E[W(y_i)]$. Then for all positive w :*

$$\text{Prob}[W(x) \geq (w + 1)E[W(x)]] < \frac{1}{e^w}$$

Both of these results are proved in the more general setting where the distributions of the h_i (and hence W) may depend on more than just the size of the input, that is we consider recurrences:

$$W(z) = a(\text{size}(z)) + \sum_{i=1}^n W(h_i(z)) \quad (4)$$

where $\text{size}(z)$ gives the size of input z .

Despite the generality of the above results, Karp notes that the assumption $E[W(x)] \geq \sum_{i=1}^n E[W(y_i)]$ in Theorem 2 is not satisfied in some important cases. Even when it is satisfied, it may be difficult to show since we need to know more than just the asymptotic value of $E[W(x)]$. Moreover, Karp mentions the problem of handling the span-recurrences of the form in Equation 2 but leaves it open. Although subsequent work [5] has presented an alternative to Theorem 1 that has some weaker assumptions, to our knowledge, no follow-up has satisfactorily extended these results to span-recurrences or weakened the assumption needed for Theorem 2.

In this paper we present a generalization of Theorem 1 to handle both span and work-recurrences for the general case of $n \geq 1$, without the restrictive assumption needed in Theorem 2. There are two key aspects of our results that distinguish them:

- Rather than having a function m such that $E[h(x)] \leq m(x)$, we require $E[\max_{i=1}^n h_i(x)] \leq m(x)$. That is, $m(x)$ must bound the expectation of the *maximum* of the problem sizes. In many applications, bounding this is not significantly more difficult than the bound needed in Theorem 1.
- Instead of checking that for all (y_1, \dots, y_n) in the support of the distribution $(h_1(x), \dots, h_n(x))$, $E[W(x)] \geq \sum_{i=1}^n E[W(y_i)]$, our theorems are parameterized by functions g_1 and g_2 for which we require $g_1(x) \geq \sum_{i=1}^n g_1(y_i)$ and $g_2(x) \geq \sum_{i=1}^n g_2(y_i)$. Often, it is easier to exhibit g_1 and g_2 with these properties than to check that the expected value has them.

Simplified forms of our main results are:

Theorem 3. *Suppose $S(x) = b(x) + \max_{i=1}^n S(h_i(x))$, where for all x , there exists some k such that the recurrence terminates after at most k recursive calls. Let g_1 be a monotone function such that for all (y_1, \dots, y_n) in the support of the distribution $(h_1(x), \dots, h_n(x))$, $\sum g_1(y_i) \leq g_1(x)$. Furthermore, assume $E[\max_{i=1}^n h_i(x)] \leq m(x)$ where $0 \leq m(x) \leq x$ and $m(x)/x$ is non-decreasing. Assume u is a solution to the recurrence $u(x) = b(x) + u(m(x))$. If b and u are continuous and b is strictly increasing on the set $\{x \mid b(x) > 0\}$, then for all positive integers w and x such that $g_1(x) \geq 1$,*

$$\text{Prob}[S(x) > u(x) + wb(x)] \leq g_1(x) \left(\frac{m(x)}{x}\right)^w$$

Theorem 4. *Suppose $W(x) = a(x) + \sum_{i=1}^n W(h_i(x))$, and the same assumptions about termination of the recurrence and $m(x)$ hold as in Theorem 3. Let g_1 and g_2 be functions such that for all (y_1, \dots, y_n) in the support of the distribution $(h_1(x), \dots, h_n(x))$, $\sum g_1(y_i) \leq g_1(x)$ and $\sum g_2(y_i) \leq g_2(x)$. Now, let $u(x)$ be a solution to the recurrence $u(x) = a(x)/g_2(x) + u(m(x))$. If a/g_2 and u are continuous and a/g_2 is strictly increasing on the set $\{x \mid a(x)/g_2(x) > 0\}$, then for all positive integers w and $g_1(x) \geq 1$:*

$$\text{Prob}[W(x) > g_2(x)u(x) + wa(x)] \leq g_1(x) \left(\frac{m(x)}{x}\right)^w$$

Notice that the tail bound has an additional factor of $g_1(x)$ compared with Theorem 1. Generally we will use $g_1(x) = x$, and we will be interested in w such that the term $(m(x)/x)^w$ is already exponentially small compared to $g_1(x)$, so we will still be able to show that the desired bound holds with high probability. Note that Theorem 1 is indeed a special case of these results, since in the degenerate case where $n = 1$, we may take $g_1(x) = g_2(x) = 1$, and we obtain Theorem 1. Moreover, we prove a more general form that handles the case when the distribution of the h_i may depend on more than just the problem size.

Our proofs are based on the proof of Theorem 1, so we begin by outlining Karp's proof of Theorem 1 in a more general setting than the version quoted above (§2). We focus only on those parts that will subsequently

change in our extensions. Next, we show how to extend the proof to handle span recurrences (§3) to obtain Theorem 3. Then, we show that in certain cases, bounding a work recurrence can be reduced to bounding span recurrences (§4), so that we can deduce Theorem 4 from Theorem 3. We then apply our method to some examples (§5). Finally, we discuss related work (§6).

2 Karp's method

In this section we give Karp's proof of a more general version of Theorem 1, where W and h are functions of problem instances rather than just their sizes. Let I be the domain of problem inputs, and let $size : I \rightarrow \mathbb{R}^+$ give the size of a problem. Then, we are interested in a recurrence of the form:

$$W(z) = a(size(z)) + W(h(z)) \quad (5)$$

where $h(z)$ is a random variable such that $0 \leq size(h(z)) \leq size(z)$. Let us first clarify what we mean by such a probabilistic recurrence. We do not mean that $W(z)$, as a function from the sample space Ω to \mathbb{R} , is equal to the expression on the right; rather, this is a statement about equality of *distributions*. That is, for all z and $r \in \mathbb{R}$:

$$Prob[W(z) > r] = \sum_y Prob[h(z) = y] \cdot Prob[W(y) > r - a(size(z))] \quad (6)$$

Since we are interested in bounding $Prob[W(z) > r]$, the recurrence only needs to be an inequality:

$$Prob[W(z) > r] \leq \sum_y Prob[h(z) = y] \cdot Prob[W(y) > r - a(size(z))] \quad (7)$$

The proof relies on the following lemma:

Lemma 1 ([12, Lemma 3.1]). *Let X be a random variable with values in the range $[0, x]$. Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is a non-negative function such that $f(0) = 0$, and there exists some constant c such that for all $y \geq c$, $f(y) = 1$ and $f(y)/y$ is non-decreasing on the interval $(0, c]$. Then:*

$$E[f(X)] \leq \frac{E[X] f(\min(x, c))}{\min(x, c)}$$

Proof. We start by showing that for all y such that $0 \leq y \leq x$,

$$f(y) \leq \frac{yf(\min(x, c))}{\min(x, c)}$$

If $y = 0$ the result is immediate since $f(0) = 0$. Otherwise, if $x \leq c$, then we have $f(y)/y \leq f(x)/x$, and the result follows. For the case where $x > c$, we have two sub-cases:

- If $y \leq c$, then we again have $f(y)/y \leq f(c)/c$ and so we are done.
- If $y > c$, then $f(y) = 1 = f(c) \leq f(c) \frac{y}{c}$.

Letting F be the CDF of X , we have:

$$E[f(X)] = \int_0^x f(y) dF(y) \leq \int_0^x \frac{yf(\min(x, c))}{\min(x, c)} dF(y) = \frac{E[X] f(\min(x, c))}{\min(x, c)} \quad \square$$

Theorem 5 ([12, Theorem 3.3]). *Let W satisfy the recurrence in Equation 7 for z such that $size(z) > d$. That is, there is some point d at which the recurrence ends. Assume $a(x) = 0$ for $x \in [0, d]$ and $a(x)$ is monotone and continuous on (d, ∞) . Suppose $E[h(x)] \leq m(x)$ where $0 \leq m(x) \leq x$ and $m(x)/x$ is non-decreasing. Assume $u : \mathbb{R} \rightarrow \mathbb{R}$ is a monotone solution to the recurrence $u(x) \geq a(x) + u(m(x))$ which is continuous on (d, ∞) . Let $u' : \mathbb{R} \rightarrow \mathbb{R}^+$ be an inverse of u above d , that is: $u'(u(x)) = x$ for $x > d$, and similarly $u(u'(y)) = y$ for $y > u(d)$. Assume for all $x \leq d$, $u(x) = u(d)$, and $u(x)$ bounds the support of $W(x)$ for $x \leq d$.*

Then for all z and r , $Prob[W(z) > r] \leq D_r(size(z))$, where:

1. If $r \leq u(d)$ then $D_r(x) = 1$
2. If $r > u(d)$:
 - (a) If $x \leq d$ then $D_r(x) = 0$
 - (b) If $x > d$ and $u(x) \geq r$ then $D_r(x) = 1$
 - (c) If $x > d$ and $u(x) < r$ then

$$D_r(x) = \left(\frac{m(x)}{x} \right)^{\left\lceil \frac{r-u(x)}{a(x)} \right\rceil} \frac{x}{u'(r - a(x) \left\lceil \frac{r-u(x)}{a(x)} \right\rceil)}$$

When $r = u(x) + wa(x)$ for integer w , $D_r(x)$ simplifies to the bound stated in Theorem 1 in the introduction.

Note: The statement here is slightly different than Theorem 3.3 in [12]. In particular, there u is a minimal, invertible solution to the deterministic recurrence and the result uses the inverse u^{-1} rather than some u' . Furthermore, a is assumed to be strictly increasing on (d, ∞) and continuous everywhere, and m is also continuous (so as to imply continuity and invertibility of u). This continuity and strictly increasing assumption rules out the case where $a(x) = 1$ for $x > d$ and $a(x) = 0$ for $x \leq d$, which arises in some examples, so Karp has a separate result for this case. However, the proof does not actually rely on continuity of a everywhere or that it be strictly increasing instead of merely monotone, so we prefer the formulation above which avoids the need for an additional separate theorem.

Proof. We will only give the proof in the case where for all z , there exists some k such that $\text{Prob}[\text{size}(h^k(z)) > d] = 0$. See [15] for discussion of the proof in a more general setting.

For concision, let us define $K_r(z) = \text{Prob}[W(z) > r]$. The strategy of the proof is to inductively define a sequence of functions K_r^i for $i \in \mathbb{N}$, then show that $K_r(z)$ is bounded by $\sup_i K_r^i(z)$, and finally prove by induction on i that D bounds each of the K_r^i .

First, Equation 7 implies $K_r(z) \leq E[K_{r-a(\text{size}(z))}(h(z))]$ when $\text{size}(z) > d$. We will use a similar ‘‘recurrence’’ to define K_r^i by:

$$K_r^0(z) = \begin{cases} 1 & \text{if } r < u(d) \\ 0 & \text{otherwise} \end{cases}$$

$$K_r^{i+1}(z) = E[K_{r-a(\text{size}(z))}^i(h(z))]$$

For all i , $K_r^i(z) \leq K_r^{i+1}(z)$ and $K_r^i(z) \leq 1$ so $\sup_i K_r^i(z)$ exists. First we note that if $\text{size}(z) \leq d$, then $K_r(z) \leq K_r^i(z)$.

Next, we claim that if $K_r^i(z) < K_r(z)$ then $\text{Prob}[\text{size}(h^i(z)) > d] > 0$. The proof is by induction on i . If $i = 0$, then we must show $\text{Prob}[\text{size}(z) > d] > 0$, but this is immediate since if not we know $K_r(z) \leq K_r^0(z)$, contradicting our assumption of the opposite inequality. For the inductive case, $K_r^{i+1}(z) < K_r(z)$ implies $E[K_{r-a(\text{size}(z))}^i(h(z))] < E[K_{r-a(\text{size}(z))}(h(z))]$. Hence for some y such that $\text{Prob}[h(z) = y] > 0$, we must have $K_{r-a(\text{size}(z))}^i(y) < K_{r-a(\text{size}(z))}(y)$. By the induction hypothesis, we have that $\text{Prob}[\text{size}(h^i(y)) > d] > 0$, hence $\text{Prob}[\text{size}(h^{i+1}(z)) > d] \geq \text{Prob}[h(z) = y] \cdot \text{Prob}[\text{size}(h^i(y)) > d] > 0$.

This implies that $K_r(z) \leq \sup_i K_r^i(z)$, for if not, then for all i we would have $K_r^i(z) < K_r(z)$, which would imply that for all i there would be some non-zero probability that the recursion would continue for at least i calls, contradicting our hypothesis about W .

We then show by induction on i that $K_r^i(z) \leq D_r(\text{size}(z))$. A key part of the induction involves applying Lemma 1 to the function D_r for some particular r . To do so, one must show that $D_r(x)/x$ is non-decreasing on the interval $(0, u'(r)]$. We will not reproduce Karp’s proof of this fact, since our later proofs will re-use this same D_r function. Nevertheless, the idea is to first divide up $(0, u'(r)]$ into subintervals on which $\left\lceil \frac{r-u(x)}{a(x)} \right\rceil$ is constant and show that $D_r(x)/x$ is non-decreasing on these intervals. Then, one shows that $D_r(x)/x$ is continuous at the ends of these subintervals, which proves monotonicity on $(0, u'(r)]$. This is where the continuity assumptions on a and u are used.

From there, the base case of the induction showing $K_r^i(z) \leq D_r(\text{size}(z))$ is straight-forward. For the inductive case, we have

$$\begin{aligned}
K_r^{i+1}(z) &= E \left[K_{r-a(\text{size}(z))}^i(h(z)) \right] \\
&\leq E \left[D_{r-a(\text{size}(z))}(h(z)) \right] \\
&\leq E \left[\text{size}(h(z)) \frac{D_{r-a(\text{size}(z))}(\min(\text{size}(z), u'(r - a(\text{size}(z)))))}{\min(\text{size}(z), u'(r - a(\text{size}(z))))} \right] \\
&\text{(by Lemma 1 with } X = \text{size}(h(z)), f = D_{r-a(\text{size}(z))}, \text{ and } c = u'(r - a(\text{size}(z)))) \\
&\leq m(\text{size}(z)) \frac{D_{r-a(\text{size}(z))}(\min(\text{size}(z), u'(r - a(\text{size}(z)))))}{\min(\text{size}(z), u'(r - a(\text{size}(z))))} \\
&= D_r(\text{size}(z))
\end{aligned}$$

□

3 Span Recurrences

We now extend the result from the previous section to handle probabilistic span recurrences of the form:

$$S(z) \leq a(\text{size}(z)) + \max_{i=1}^n S(h_i(z))$$

where $0 \leq \text{size}(h_i(z)) \leq \text{size}(z)$. In terms of the CDF of $S(z)$ this is:

$$\text{Prob}[S(z) > r] \leq \sum_{y_1, \dots, y_n} \text{Prob}[(h_1(z), \dots, h_n(z)) = (y_1, \dots, y_n)] \cdot \text{Prob} \left[\max_{i=1}^n S(y_i) > r - a(\text{size}(z)) \right] \quad (8)$$

Recall from the introduction that the key idea of our result is that we now require a bound on the expected *maximum* of the sizes of the subproblems $h_1(z), \dots, h_n(z)$, and we introduce a parameter g_1 such that $g_1(\text{size}(x)) \geq \sum_{i=1}^n g_1(\text{size}(y_i))$ for all (y_1, \dots, y_n) in the support of the joint distribution $(h_1(z), \dots, h_n(z))$. We start by proving a generalization of Lemma 1 which suggests the role that these assumptions will play:

Lemma 2. *Let $X : \Omega \rightarrow [0, x]^n$ be a random vector where $n \in \mathbb{N}^+$. Let $g : \mathbb{R} \rightarrow \mathbb{R}^+$ be a function such that for all (y_1, \dots, y_n) in the support of X , $0 \leq g(y_i)$ and $\sum_i g(y_i) \leq g(x)$. Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is a non-negative function such that $f(0) = 0$, and there exists some constant c such that for all $y \geq c$, $f(y) = 1$ and $f(y)/y$ is non-decreasing on the interval $(0, c]$. Then:*

$$E \left[\sum_{i=1}^n g(X_i) f(X_i) \right] \leq g(x) \frac{E[\max_{i=1}^n X_i] f(\min(x, c))}{\min(x, c)}$$

where X_i is the i th component of the random vector X .

Proof. It suffices to show that for all (y_1, \dots, y_n) in the support of X ,

$$\sum_{i=1}^n g(y_i) f(y_i) \leq g(x) \frac{(\max_{i=1}^n y_i) f(\min(x, c))}{\min(x, c)}$$

We have:

$$\sum_{i=1}^n g(y_i) f(y_i) \leq \left(\max_{j=1}^n f(y_j) \right) \left(\sum_{i=1}^n g(y_i) \right) \leq \left(\max_{j=1}^n f(y_j) \right) g(x)$$

Monotonicity of $f(x)/x$ implies f is monotone. Furthermore, by the argument in Lemma 1, we know that for each j , $f(y_j) \leq \frac{y_j f(\min(x, c))}{\min(x, c)}$. Hence, we have:

$$\begin{aligned}
\sum_{i=1}^n g(y_i) f(y_i) &\leq \left(\max_{j=1}^n f(y_j) \right) g(x) \\
&\leq f\left(\max_{j=1}^n y_j\right) g(x) \\
&\leq \frac{\left(\max_{j=1}^n y_j\right) f(\min(x, c))}{\min(x, c)} g(x)
\end{aligned}$$

□

The proof of the next result has the same overall structure as that of Theorem 5, except we use the above lemma in place of Lemma 1.

Theorem 6. *Let S satisfy the recurrence in Equation 8 for z such that $\text{size}(z) > d$. Assume for all z there exists k such that for all i_1, \dots, i_k we have $\text{Prob}[\text{size}(h_{i_1}(\dots(h_{i_k}(z)))) > d] = 0$. Assume $a(x) = 0$ for $x \in [0, d]$ and $a(x)$ is monotone and continuous on (d, ∞) . Suppose $E[\max_{i=1}^n h_i(x)] \leq m(x)$ where $0 \leq m(x) \leq x$ and $m(x)/x$ is non-decreasing. Let u, u' be as in Theorem 5, under analogous assumptions. Furthermore, let $g_1 : \mathbb{R} \rightarrow \mathbb{R}^+$ be a monotone function such that for all $x > d$, $g_1(x) \geq 1$. Assume for all z such that $\text{size}(z) > d$, and all (y_1, \dots, y_n) in the joint distribution of $(h_1(z), \dots, h_n(z))$:*

$$\sum_{i=1}^n g_1(\text{size}(y_i)) \leq g_1(\text{size}(z))$$

Then for all z and r such that $g_1(\text{size}(z)) \geq 1$, we have $\text{Prob}[S(z) > r] \leq g_1(\text{size}(z)) \cdot D_r(\text{size}(z))$ where $D_r(\text{size}(z))$ is as in Theorem 5.

Proof. Set $K_r(z) = \text{Prob}[S(z) > r]$. By the union bound, Equation 8 implies

$$K_r(z) \leq \sum_{i=1}^n E[K_{r-a(\text{size}(z))}(h_i(z))]$$

As before, we will inductively define a family of functions K^i using this same recurrence, show that their suprema bounds K , and then inductively bound the K^i . Define $K_r^i(z)$ by:

$$\begin{aligned}
K_r^0(z) &= \begin{cases} 1 & \text{if } r < u(d) \\ 0 & \text{otherwise} \end{cases} \\
K_r^{i+1}(z) &= \min\left(1, \sum_{j=1}^n E\left[K_{r-a(\text{size}(z))}^i(h_j(z))\right]\right)
\end{aligned}$$

For all i , $K_r^i(z) \leq K_r^{i+1}(z)$ and $K_r^i(z) \leq 1$, so $\sup_i K_r^i(z)$ exists. As before, if $\text{size}(z) \leq d$, then $K_r(z) \leq K_r^i(z)$. Here, if $K_r^n(z) < K_r(z)$ then $\text{Prob}[\text{size}(h_{i_1}(\dots(h_{i_n}(z)))) > d] > 0$ for some sequence i_1, \dots, i_n . The proof is by induction on n , where the base case is immediate. For the inductive case, if $K_r^{n+1}(z) < K_r(z)$ we must have that $K_r^{n+1}(z) < 1$, so $K_r^{n+1}(z) = \sum_j E[K_{r-a(\text{size}(z))}^n(h_j(z))] < \sum_j E[K_{r-a(\text{size}(z))}(h_j(z))]$. Hence for some i and y such that $\text{Prob}[h_i(z) = y] > 0$, we must have $K_{r-a(\text{size}(z))}^n(y) < K_{r-a(\text{size}(z))}(y)$. By the induction hypothesis, $\text{Prob}[\text{size}(h_{i_1}(\dots(h_{i_n}(y)))) > d] > 0$. Then the sequence i_1, \dots, i_n, i suffices. It follows that $K_r(z) \leq \sup_i K_r^i(z)$, or else our assumption about the termination of the recurrence would be violated.

When $r < u(d)$ and $\text{size}(z) \geq d$, we have $K_r^i(z) \leq 1 \leq g_1(\text{size}(z)) \cdot D_r(\text{size}(z))$. For the case when $r \geq u(d)$ we prove by induction on i that $K_r^i(z) \leq g_1(\text{size}(z)) \cdot D_r(\text{size}(z))$. The base case is similar to that

of Theorem 5. For the inductive step, the only non-trivial sub-case is when $\text{size}(z) > d$ and $r > u(\text{size}(z))$. Then we have:

$$\begin{aligned}
K_r^{i+1}(z) &\leq \sum_{j=1}^n E \left[K_{r-a(\text{size}(z))}^i(h_j(z)) \right] \\
&\leq \sum_{j=1}^n E \left[g_1(\text{size}(h_j(z))) \cdot D_{r-a(\text{size}(z))}(h_j(z)) \right] \\
&\leq g_1(\text{size}(z)) \cdot E \left[\max_{j=1}^n (\text{size}(h_j(z))) \right] \frac{D_{r-a(\text{size}(z))}(\min(\text{size}(z), u'(r - a(\text{size}(z)))))}{\min(\text{size}(z), u'(r - a(\text{size}(z))))} \\
&\text{(by Lemma 2 with } X = (\text{size}(h_1(z)), \dots, \text{size}(h_n(z))), f = D_{r-a(\text{size}(z))}, \\
&\quad g = g_1, \text{ and } c = u'(r - a(\text{size}(z)))) \\
&\leq g_1(\text{size}(z)) \cdot m(\text{size}(z)) \frac{D_{r-a(\text{size}(z))}(\min(\text{size}(z), u'(r - a(\text{size}(z)))))}{\min(\text{size}(z), u'(r - a(\text{size}(z))))} \\
&= g_1(\text{size}(z)) \cdot D_r(\text{size}(z))
\end{aligned}$$

In the second line, when we apply the induction hypothesis, we must check that $r - a(\text{size}(z)) \geq u(d)$. This follows from the fact that $r > u(\text{size}(z)) \geq a(\text{size}(z)) + u(m(\text{size}(z))) \geq a(\text{size}(z)) + u(d)$. \square

4 Work Recurrences

Finally we show in this section that in some cases, work-recurrences can be transformed into analogous span-like recurrences. Our starting point is a recurrence:

$$W(z) \leq a(\text{size}(z)) + \sum_{i=1}^n W(h_i(z)) \quad (9)$$

which terminates when $\text{size}(z) \leq d$. Let $g_2 : \mathbb{R} \rightarrow \mathbb{R}^+$ be a function such that for all (y_1, \dots, y_n) in the support of the distribution $(\text{size}(h_1(z)), \dots, \text{size}(h_n(z)))$, $\sum g_2(y_i) \leq g_2(x)$. Consider the random variable $H(z) = W(z)/g_2(\text{size}(z))$. We have:

$$\begin{aligned}
H(z) &\leq \frac{a(\text{size}(z))}{g_2(\text{size}(z))} + \sum_{i=1}^n \left(\frac{1}{g_2(\text{size}(z))} W(h_i(z)) \right) \\
&= \frac{a(\text{size}(z))}{g_2(\text{size}(z))} + \sum_{i=1}^n \left(\frac{g_2(\text{size}(h_i(z)))}{g_2(\text{size}(z))} H(h_i(z)) \right) \\
&\leq \frac{a(\text{size}(z))}{g_2(\text{size}(z))} + \sum_{i=1}^n \left(\frac{g_2(\text{size}(h_i(z)))}{g_2(\text{size}(z))} \max_{j=1}^n H(h_j(z)) \right) \\
&= \frac{a(\text{size}(z))}{g_2(\text{size}(z))} + \sum_{i=1}^n \left(\frac{g_2(\text{size}(h_i(z)))}{g_2(\text{size}(z))} \right) \max_{j=1}^n H(h_j(z)) \\
&\leq \frac{a(\text{size}(z))}{g_2(\text{size}(z))} + \max_{j=1}^n H(h_j(z))
\end{aligned}$$

That is, we see H obeys a span-like recurrence with toll-function a/g_2 . Since $\text{Prob}[H(z) > r] = \text{Prob}[W(z) > r g_2(\text{size}(z))]$, we can apply the theorem from the previous section to H and thereby derive a tail-bound for W , giving us:

Theorem 7. *Let W satisfy the recurrence in Equation 9 for z such that $\text{size}(z) > d$. Assume for all z there exists k such that for all i_1, \dots, i_k we have $\text{Prob}[\text{size}(h_{i_1}(\dots(h_{i_k}(z)))) > d] = 0$. Suppose $E[\max_{i=1}^n h_i(x)] \leq$*

$m(x)$ where $0 \leq m(x) \leq x$ and $m(x)/x$ is non-decreasing. Let u be a solution to the recurrence $u(x) \leq a(x)/g_2(x) + u(m(x))$ under analogous assumptions as the previous theorems. Furthermore, let $g_1, g_2 : \mathbb{R} \rightarrow \mathbb{R}^+$ be monotone functions where for all $x > d$, $g_1(x) \geq 1$. Assume for all z such that $\text{size}(z) > d$, and all (y_1, \dots, y_n) in the joint distribution of $(h_1(z), \dots, h_n(z))$:

$$\sum_{i=1}^n g_1(\text{size}(y_i)) \leq g_1(\text{size}(z)) \quad \text{and} \quad \sum_{i=1}^n g_2(\text{size}(y_i)) \leq g_2(\text{size}(z))$$

Finally, let $a(x)/g(x) = 0$ for $x \in [0, d]$ and assume $a(x)/g(x)$ is monotone and continuous on (d, ∞) . Then for all z and r such that $g_1(\text{size}(z)) \geq 1$, we have $\text{Prob}[W(z) > r] \leq g_1(\text{size}(z)) \cdot D_{r/g_2(\text{size}(z))}(\text{size}(z))$ where $D_r(\text{size}(z))$ is as in Theorem 5, using the toll function a/g_2 .

5 Examples

We now apply the previous results to examples.

Quicksort Let l be a list and write $|l|$ for the length of the list. The work and span owing to the comparisons involved in parallel randomized quicksort satisfy the following recurrences:

$$\begin{aligned} W(l) &= (|l| - 1) + W(h_1(l)) + W(h_2(l)) \\ S(l) &= \log |l| + \max(S(h_1(l)), S(h_2(l))) \end{aligned}$$

where $E[\max(|h_1(l)|, |h_2(l)|)] \leq \frac{7|l|}{8}$ when $|l| > 1$ and is 0 otherwise; moreover, $|h_1(l)| + |h_2(l)| \leq |l| - 1$. Note that this is a case where the distributions of the sizes of the sublists $h_1(l)$ and $h_2(l)$ depend on more than just the size of l : in particular, if l contains multiple versions of the same element, then the sublists will tend to be smaller than if all elements of l are distinct. Therefore, the more general formulations of Theorem 6 and Theorem 7 are useful here. For span, we apply Theorem 6 with $m(x) = \frac{7x}{8}$, $u(x) = (\log_{8/7} x + 1)^2$, and $g_1(x) = x$, to get that for all positive integers w ,

$$\text{Prob} \left[S(l) > (\log_{8/7} |l| + 1)^2 + w \log |l| \right] \leq |l| \left(\frac{7}{8} \right)^w$$

when $|l| \geq 1$. Take w to be $\lceil \frac{k}{(\log 8/7)} \log_{8/7} |l| \rceil$ for integer k so that we have

$$\text{Prob} \left[S(l) > (k+1)(\log_{8/7} |l| + 1)^2 \right] \leq \left(\frac{1}{|l|} \right)^{\frac{k}{\log \frac{8}{7}} - 1}$$

For the work, we take g_2 to be:

$$g_2(x) = \begin{cases} \frac{1}{2} & \text{if } x \leq 1 \\ 1 & \text{if } 1 < x < 2 \\ x - 1 & \text{if } x \geq 2 \end{cases}$$

In this case, in order to solve the recurrence for u , it is easier if we take m to be

$$m(x) = \begin{cases} 0 & \text{if } x < 8/7 \\ \frac{7x}{8} & \text{if } x \geq 8/7 \end{cases}$$

Applying Theorem 7 with $u(x) = \log_{8/7}(x) + 1$ we obtain:

$$\text{Prob} \left[W(l) > (|l| - 1)(\log_{8/7} |l| + 1) + w(|l| - 1) \right] \leq |l| \left(\frac{7}{8} \right)^w$$

Here taking w to be $\lceil k \log_{8/7} |l| \rceil$ for integer k we have

$$\text{Prob} \left[W(l) > (k+1) \left(|l| \log_{8/7} |l| + |l| \right) \right] \leq \left(\frac{1}{|l|} \right)^{k-1}$$

Karp shows in [12] that the quicksort work recurrence does satisfy the hypotheses of Theorem 2, which one can check knowing that $E[W(|l|)] = 2((|l|+1)(H_{|l|}-1) - (|l|-1))$, where H_n is the n th Harmonic number. However, it is much easier to check that g_2 satisfies the analogous hypothesis of our result. Moreover, the resulting bound obtained by Karp's Theorem 2,

$$\text{Prob} [W(l) > (a+1)E[W(l)]] \leq \frac{1}{e^a}$$

is weaker than the bound we obtain above.

Height of search trees Consider random binary search trees generated by successively inserting the elements of a permutation of $\{1, \dots, n\}$, with all permutations equally likely. The height $H(n)$ of such a tree obeys the following probabilistic recurrence:

$$H(n) = 1 + \max(H(h_1(n)), H(h_2(n)))$$

where again $E[\max(h_1(n), h_2(n))] \leq \frac{7n}{8}$ and $h_1(n) + h_2(n) = n - 1$. Therefore we can apply Theorem 6 with similar choice of m as in the previous examples, and with $u(n) = \log_{8/7}(n) + 1$ to get:

$$\text{Prob} \left[H(n) > (\log_{8/7} n + 1) + w \right] \leq n \left(\frac{7}{8} \right)^w$$

Again, for appropriate choice of w this yields:

$$\text{Prob} \left[H(n) > (k+1)(\log_{8/7} n + 1) \right] \leq \left(\frac{1}{n} \right)^{k-1}$$

A similar argument works for a large class of tree structures, such as radix trees and quad trees, so long as we can bound the expected value of the maximum of the sizes of the subtrees.

6 Conclusion and Related Work

We have presented a way to extend Theorem 1 to span and work recurrences, answering the open questions posed by Karp in [12]. Our result for work recurrences is easier to apply than Theorem 2 and gives stronger bounds in cases like the analysis of quicksort.

Of course, just as with the original theorems, more precise bounds can be obtained by instead using more advanced techniques. The books by Flajolet and Sedgewick [11] and Drmota [9] give an overview of techniques from analytic combinatorics and probability theory that can be used to analyze processes that give rise to these kinds of recurrences.

Nevertheless, Karp has argued convincingly that it is still useful to have easy to apply cook-book techniques for use in situations where less precise results are often obtained using ad-hoc methods. Cook-book methods are common and well known for analyzing deterministic divide-and-conquer recurrences (e.g. [1, 14, 10]). However, for randomized divide-and-conquer, there are far fewer results. The method of Roura [14] applies to the recurrence for the expected number of comparisons in quicksort. Bazzi and Mitter [2] extend the Akra-Bazzi method [1] to derive asymptotic bounds for expected work recurrences.

Chaudhuri and Dubhashi [5] study the unary probabilistic recurrences of Theorem 1 and give an alternative result that does not require $m(x)/x$ to be non-decreasing (but in this case, the bound is weaker than Theorem 1). They suggest that one advantage of their proof is that it only uses common tools from probability theory like Markov's inequality and Chernoff bounds. It would be interesting to see if some of the ideas we used here can also be used to extend their version to handle general span and work recurrences.

In a technical report, Karpinski and Zimmermann [13] also modify Theorem 1 to try to handle span and work recurrences. However, the bounds they obtain appear to be weaker than ours. For instance, applied to the span recurrence of quicksort, their result gives a bound of $(1 - \frac{1}{n})^w$ where our Theorem 6 gives $n(\frac{7}{8})^w$. That is, the bounds they give do not appear to be sufficient to establish high probability bounds on run-time. Moreover, their proof modifies the $D_r(x)$ function and applies Lemma 1, but they do not give a proof for why $D_r(x)/x$ remains non-decreasing, and Karp's strategy for showing that it is non-decreasing does not appear to generalize to their version.

Acknowledgments The author thanks Robert Harper, Jean-Baptiste Tristan, Victor Luchangco, Guy Blelloch, Guy L. Steele Jr., and Kuen-Bang Hou (Favonia) for helpful discussions. This research was conducted with U.S. Government support under and awarded by DoD, Air Force Office of Scientific Research, National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of this funding agency.

References

- [1] M. Akra and L. Bazzi. On the solution of linear recurrence equations. *Comp. Opt. and Appl.*, 10(2): 195–210, 1998. doi: 10.1023/A:1018373005182. URL <http://dx.doi.org/10.1023/A:1018373005182>.
- [2] L. Bazzi and S. K. Mitter. The solution of linear probabilistic recurrence relations. *Algorithmica*, 36(1):41–57, 2003. doi: 10.1007/s00453-002-1003-4. URL <http://dx.doi.org/10.1007/s00453-002-1003-4>.
- [3] G. Blelloch and J. Greiner. Parallelism in sequential functional languages. In *Proceedings of the 7th International Conference on Functional Programming Languages and Computer Architecture*, pages 226–237, 1995. ISBN 0-89791-719-7.
- [4] R. D. Blumofe and C. E. Leiserson. Scheduling multithreaded computations by work stealing. *J. ACM*, 46(5):720–748, 1999. doi: 10.1145/324133.324234. URL <http://doi.acm.org/10.1145/324133.324234>.
- [5] S. Chaudhuri and D. P. Dubhashi. Probabilistic recurrence relations revisited. *Theor. Comput. Sci.*, 181(1):45–56, 1997. doi: 10.1016/S0304-3975(96)00261-7. URL [http://dx.doi.org/10.1016/S0304-3975\(96\)00261-7](http://dx.doi.org/10.1016/S0304-3975(96)00261-7).
- [6] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms (3. ed.)*. MIT Press, 2009. ISBN 978-0-262-03384-8. URL <http://mitpress.mit.edu/books/introduction-algorithms>.
- [7] L. Devroye. A note on the height of binary search trees. *J. ACM*, 33(3):489–498, 1986. doi: 10.1145/5925.5930. URL <http://doi.acm.org/10.1145/5925.5930>.
- [8] M. Drmota. An analytic approach to the height of binary search trees. *Algorithmica*, 29(1):89–119, 2001. doi: 10.1007/BF02679615. URL <http://dx.doi.org/10.1007/BF02679615>.
- [9] M. Drmota. *Random Trees*. Springer-Verlag Wien, 2009. ISBN 978-3-211-75355-2.
- [10] M. Drmota and W. Szpankowski. A master theorem for discrete divide and conquer recurrences. *J. ACM*, 60(3):16:1–16:49, 2013. doi: 10.1145/2487241.2487242. URL <http://doi.acm.org/10.1145/2487241.2487242>.
- [11] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. ISBN 978-0-521-89806-5. URL <http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521898065>.
- [12] R. M. Karp. Probabilistic recurrence relations. *J. ACM*, 41(6):1136–1150, 1994. doi: 10.1145/195613.195632. URL <http://doi.acm.org/10.1145/195613.195632>.

- [13] M. Karpinski and W. Zimmermann. Probabilistic recurrence relations for parallel divide-and-conquer algorithms. Technical Report TR-91-067, International Computer Science Institute (ICSI), 1991. URL <https://www.icsi.berkeley.edu/ftp/global/pub/techreports/1991/tr-91-067.pdf>.
- [14] S. Roura. Improved master theorems for divide-and-conquer recurrences. *J. ACM*, 48(2):170-205, 2001. doi: 10.1145/375827.375837. URL <http://doi.acm.org/10.1145/375827.375837>.
- [15] N. Young. Answer to: Understanding proof of theorem 3.3 in Karp's probabilistic recurrence relations. Theoretical Computer Science Stack Exchange. URL <http://cstheory.stackexchange.com/q/37144>. URL:<http://cstheory.stackexchange.com/q/37144> (version: 2016-12-13).