Estimating Mean Completion Times of a Fork-Join Barrier Synchronization

Eric Bach
Rajesh Mansharamani
John Strikwerda

Technical Report #1216

February 1994

Estimating Mean Completion Times of a Fork-Join Barrier Synchronization†

Eric Bach, Rajesh Mansharamani, and John Strikwerda

Computer Sciences Department
1210 W. Dayton St.
Madison, WI 53706.

Tel: (608) 262-7997
Email: {bach, mansha, strik}@cs.wisc.edu

February 21, 1994

Abstract. In simulation studies of parallel processors, it is useful to consider the following abstraction of a parallel program. A job is partitioned into n processes, whose running times are independent random variables $X_1, \ldots, X_n$. As a measure of performance we consider the normalized job completion time $S = \max\{X_i\} / \sum_{i=1}^{n} X_i$. We consider a simple approximation to the expected value of S, valid asymptotically whenever the X_i's are bounded, and assess its accuracy as a function of n both theoretically and experimentally. The approximation is easy to compute and involves only the first two moments of X_i.

1 Introduction

In this paper we study a simple performance metric for parallel programs. We are interested in so-called fork-join programs, of the following type. A job splits into n processes that run independently, then wait for the last one to complete. As a measure of performance we consider the ratio

$$S := \frac{\max\{T_i\}}{D}, \tag{1}$$

where T_i is the time taken by process i, and $D = \sum_{i=1}^{n} T_i$ is the total demand. Roughly, this ratio tells us how fast one can solve a problem in parallel, relative to the cost of solving it sequentially. Thus it is the inverse of the usual "speedup" studied in parallel programming. It can also be thought of as a job completion time, normalized by demand.

This ratio, assuming process times are i.i.d. samples from a uniform distribution, has been used in simulation studies of multiprogrammed parallel processor scheduling policies [9, 11]. In such studies it is useful to know the expected value of S, which is proportional to the mean completion time of a simulated

†This research was partially supported by the National Science Foundation under grants DCR-9208639, DMS-9208049 and CCR-9024144

jobs. In [13] two expressions for

$$\overline{S} := E\left[\frac{\max\{T_i\}}{D}\right] \tag{2}$$

are given: a closed form (not usable beyond $n = 50$ due to cancellation) and an infinite series, which leads to an $O(n^2)$ algorithm to approximate $\overline{S}$. Both of these results assume process times are uniformly distributed.

In this paper we propose a simple two-moment approximation for $\overline{S}$ under more general assumptions for process service times than [13]. More specifically, we assume that the total demand D is partitioned into processes of length

$$T_j = \frac{X_j}{\sum_{i=1}^{n} X_i} D, \tag{3}$$

where $X_1, \ldots, X_n$ are i.i.d. bounded positive random variables, independent of D. We use the Central Limit Theorem to estimate $\overline{S}$ as a function of n when $n \to \infty$, thereby obtaining approximations useful for large n. We validate our approximation against simulation for several job length distributions.

The results of this paper extend previous work in several ways. Several authors have studied the distributions of S and $1/S$ under various assumptions about the parent distribution. (See [1, 2, 5, 6] and references therein.) The emphasis in these works was on asymptotic results for unbounded random variables, in part because $1/S$ is asymptotically normal otherwise. (Unfortunately, this doesn't tell us much about the expected value of S: when Z is normal, $E[1/Z]$ fails to exist.) In contrast, we are interested in numerical estimates for $\overline{S}$ and in formulas that relate the expected performance of fork-join programs to simple system parameters, such as mean and variance of process length. Such formulas are important because detailed distributional information is rarely available in a real situation. In the uniform case, our approximation leads to an asymptotic series for $\overline{S}$. As this series is more accurate as n increases, this complements the methods of [13].

Although we are primarily concerned with bounded random variables, it it is worthwhile to remark that $E[S]$ can be estimated rather precisely in the important unbounded case where $X_1, \ldots, X_n$ are i.i.d. samples from a $\Gamma(k)$ distribution. This distribution is often used as a model for process times; for example, the exponential distribution has $k = 1$. In this case, one could use the asymptotic distribution of $1/S$ given by Darling [6] to estimate $E[S]$; we discuss an alternative approach below.

The remainder of this paper is organized as follows. Section 2 derives the approximation for $\overline{S}$ in terms of the first two moments of X. Section 3 provides validations of the approximation against simulation estimates of $\overline{S}$.

2 Approximations for $\overline{S}$

In this section we derive an approximation for $\overline{S}$, as defined by (2) and (3), in terms of n and the first two moments of X. Our approximation $\hat{S}$ is asymptotically exact, in the sense that $\lim_{n\to\infty} \overline{S}/\hat{S} = 1$. As we are concerned only with ratios of completion times we will ignore total demand and without loss of generality set $D = 1$ in the remainder of this paper.

The idea behind our approximation is as follows. We have

$$S = \frac{\max\{X_i\}}{\sum_{i=1}^{n} X_i} \tag{4}$$

If the X_i's are i.i.d. and n is large, then any particular X_i will contribute very little to the sum and we incur little error in replacing one of them, say X_n, by its maximum value M. On the other hand, M will be a good approximation to $\max\{X_i\}$, since it is unlikely that all of the X_i's are small. Therefore the mean value of S should be close to

$$\hat{S} := E\left[\frac{M}{\sum_{i=1}^{n-1} X_i + M}\right]. \tag{5}$$

By the law of large numbers, we have

$$\hat{S} \sim \frac{1}{1 + (n-1)E[X]/M},$$

and we can refine this by applying the central limit theorem to $\sum_{i=1}^{n-1} X_i$. (Technically we should use $E[X_i]$ rather than $E[X]$, but our convention will be to drop such subscripts.)

We now carry out this program precisely. We will assume that $0 \leq X \leq M$ and that M is the essential supremum of X. (That is, no random variable agreeing with X almost everywhere has a smaller bound.) Thus all moments of X will exist; to avoid degenerate cases we assume that the mean and variance of X are positive. We first prove that we can replace $\overline{S}$ by $\hat{S}$ asymptotically.

Proposition 2.1 *We have* $\lim_{n\to\infty} \overline{S}/\hat{S} = 1$.

Proof. We first prove a lower bound. Choose $\epsilon > 0$. Since $X_n \leq M$, we have

$$E\left[\frac{\max\{X_i\}}{\sum_{i=1}^{n} X_i}\right] \geq E\left[\frac{\max\{X_i\}}{M} \frac{M}{\sum_{i=1}^{n-1} X_i + M}\right]$$

Letting ν be the probability measure, this is at least

$$(1-\epsilon)\int_{\max\{X_i\}>(1-\epsilon)M} \frac{M}{\sum_{i=1}^{n-1} X_i + M} d\nu = (1-\epsilon)\left(\hat{S} - \int_{\max\{X_i\}\leq(1-\epsilon)M} \frac{M}{\sum_{i=1}^{n-1} X_i + M} d\nu\right).$$

The second integral is bounded by α^n, where $\alpha = \Pr[X \leq (1-\epsilon)M]$. (Note that $\alpha < 1$.) Since $\hat{S} \geq 1/n$, we have $\overline{S} \geq (1-\epsilon)(1-n\alpha^n)\hat{S}$.

We now prove an upper bound. Since $\max\{X_i\} \leq M$ and $X_n \geq 0$, we have

$$\frac{\max\{X_i\}}{\sum_{i=1}^{n} X_i} \leq \frac{M}{\sum_{i=1}^{n-1} X_i + M}(1 + \frac{M}{\sum_{i=1}^{n-1} X_i}). \tag{6}$$

Choose any c satisfying $0 < c < 1$. We will call $\sum_{i=1}^{n-1} X_i$ "small" if $\sum_{i=1}^{n-1} X_i \leq c(n-1)E[X]$, and "large" otherwise. From large deviation theory (see Appendix), there is a $\beta < 1$ for which $\Pr[\sum_{i=1}^{n-1} X_i \text{ small}] \leq \beta^{n-1}$. We may in fact take

$$\beta = \min_{\lambda \geq 0}\left\{\Psi_{X-cE[X]}(-\lambda)\right\}, \tag{7}$$

where Ψ denotes the moment generating function. We now express $\overline{S}$ as

$$E\left[S|\ \textstyle\sum_{i=1}^{n-1} X_i \text{ small}\right] \Pr[\ \textstyle\sum_{i=1}^{n-1} X_i \text{ small}\ ] + E\left[S|\textstyle\sum_{i=1}^{n-1} X_i \text{ large}\right] \Pr[\ \textstyle\sum_{i=1}^{n-1} X_i \text{ large}\ ].$$

The first term is at most β^{n-1}, since $S \leq 1$. Since $E[T|A]\Pr[A] \leq E[T]$ whenever $T \geq 0$, we can use (6) to bound the second term by

$$\hat{S}(1 + \frac{1}{c(n-1)E[X]}).$$

Since $\hat{S} \geq 1/n$, this proves $\overline{S}/\hat{S} \leq 1 + n\beta^n + (c(n-1)E[X])^{-1} = 1 + O(1/n)$. ∎

Given sufficient information about the distribution of X, the bounds in this proof can be used to find numerical estimates for $\overline{S}/\hat{S}$. Alternatively, one can use the following quick and dirty universal bounds.

Proposition 2.2 *We have*

$$\frac{E[X]}{M} \leq \frac{E[\max\{X_i\}]}{M} \leq \frac{\overline{S}}{\hat{S}} \leq 2.$$

Proof. To prove the upper bound, observe that

$$\frac{\overline{S}}{\hat{S}} = \frac{\max\{X_i\}}{M} \frac{\sum_{i=1}^{n-1} X_i + M}{\sum_{i=1}^{n-1} X_i + X_n} = \frac{\max\{X_i\}}{\sum_{i=1}^{n-1} X_i + X_n} \frac{\sum_{i=1}^{n-1} X_i + M}{M},$$

and estimate the first expression if $\sum_{i=1}^{n-1} X_i \geq M$, the second otherwise.

To prove the lower bound, let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics of the X_i's. We have

$$E\left[\frac{\max\{X_i\}}{\sum_{i=1}^{n} X_i}\right] \geq E\left[\frac{\max\{X_i\}}{M} \cdot E\Big[\frac{M}{M + \sum_{i=1}^{n-1} X_{(i)}} \mid \max\{X_i\}\Big]\right].$$

Given any value of $\max\{X_i\}$, $X_{(1)} + \ldots + X_{(n-1)}$ is stochastically bounded by the unconditional random variable $X_1 + \ldots X_{n-1}$, i so

$$E\Big[\frac{M}{M + \sum_{i=1}^{n-1} X_{(i)}} \mid \max\{X_i\}\Big] \geq E\Big[\frac{M}{M + \sum_{i=1}^{n-1} X_i}\Big].$$

This proves $\overline{S}/\hat{S} \geq E[\max\{X_i\}]/M$; clearly this is at least $E[X]/M$. ∎

We now explain how to obtain numerical estimates for $\hat{S}$. If we let $V_i = X_i/M$, we have

$$\hat{S} = E\left[\frac{1}{1 + \sum_{i=1}^{n-1} V_i}\right] = \frac{1}{1 + (n-1)E[V]} E\left[\frac{1}{1 + Z_n}\right], \tag{8}$$

where

$$Z_n := \frac{\sum_{i=1}^{n-1}(V_i - E[V])}{1 + (n-1)E[V]}.$$

Since the V_i's are i.i.d. with $0 \leq V_i \leq 1$, we expect Z_n to be small, comparable to $1/\sqrt{n}$, and this suggests using a Taylor series for $1/(1 + Z_n)$. These ideas lead to an asymptotic series for $\hat{S}$, in terms of the moments of X. Below we give the two-moment version.

Proposition 2.3 *Let $C_X = \sigma_X/E[X]$ be the coefficient of variation of X. As $n \to \infty$, we have*

$$\hat{S} = \frac{1}{1 + (n-1)E[X]/M} \left(1 + \frac{C_X^2}{n-1} + O(\frac{1}{n^2})\right).$$

Proof. From a finite Taylor expansion, we have

$$E\left[\frac{1}{1+Z_n}\right] = 1 - E[Z_n] + E[Z_n^2] - E[Z_n^3] + E\left[\frac{Z_n^4}{1+Z_n}\right]. \tag{9}$$

Let $m = n-1$; then we have

$$Z_n = \frac{1}{E[V]}(1+\frac{1}{mE[V]})^{-1} \times \frac{\sum_{i=1}^m (V_i - E[V_i])}{m}.$$

We recognize the second factor as a sample mean, whose moments are known from sampling theory (see Appendix). Therefore

$$E[Z_n] = 0; \tag{10}$$

$$E[Z_n^2] = \frac{1}{E[V]^2}(1+\frac{1}{mE[V]})^{-2}\frac{\sigma_V^2}{m} = \frac{C_V^2}{m} + O(\frac{1}{m^2}); \tag{11}$$

$$E[Z_n^3] = O(\frac{1}{m^2}). \tag{12}$$

(Here and below, the implied constants depend on the moments of X.) We now prove a sharp bound on the expected value of $Z_n^4/(1+Z_n)$. Observe that

$$Z_n \le \frac{m(1-E[V])}{1+mE[V]} = O(1)$$

and

$$\frac{1}{1+Z_n} = \frac{1+mE[V]}{1+\sum_{i=1}^m V_i} = O(m).$$

From large deviation theory (see Appendix), there is a constant γ with $0 < \gamma < 1$ such that $\Pr[\sum_{i=1}^m V_i \le mE[V]/2] \le \gamma^m$, and this leads to

$$E\left[\frac{Z_n^4}{1+Z_n}\right] \le 2E[Z_n^4] + O(m\gamma^m). \tag{13}$$

We bound the fourth moment of Z_n using the central limit theorem. We have

$$Z_n = \frac{\sigma_V}{\sqrt{m}E[V]}(1+\frac{1}{mE[V]})^{-1} \times \frac{\sum_{i=1}^m (V_i - E[V_i])}{\sigma_V\sqrt{m}},$$

and the second factor converges in distribution to a standard normal random variable. In our case, all moments of the second factor converge to the corresponding moments of a $N(0,1)$ random variable (see Appendix), and we have

$$E[Z_n^4] = O(\frac{1}{m^2}). \tag{14}$$

We substitute (9) in (8) and apply (10)–(14) to get the result. ∎

Using Propositions 2.1 and 2.3, we have the approximation

$$\overline{S} \sim \frac{1}{1+(n-1)E[X]/M}\left(1+\frac{C_X^2}{n-1}+O(\frac{1}{n^2})\right). \tag{15}$$

which uses only the first two moments of X. Further approximations, using higher moments, could be computed by a straightforward extension to the proof of Proposition 2.3.

For the uniform case (X_i i.i.d. $U(0,1)$), our method gives an asymptotic series for $\overline{S}$. We observe that S has the same distribution as $1/(1+\sum_{i=1}^{n-1} X_i)$, as is easily proved by conditioning on the maximum. (Compare [6, page 451].) Therefore,

$$\overline{S} = \hat{S} = \frac{2}{n+1}\left[1+\frac{1}{3n}-\frac{2}{3n^2}+\frac{4}{45n^3}+\ldots\right].$$

On the other hand, one should not expect (15) to converge quickly for distributions with rapidly vanishing tails. Consider, as a limiting case, X_i's that are i.i.d. $\Gamma(k)$ random variables. It is of independent interest to compute $E[S]$ in this case, and this can be done as follows. Since each X_i can be represented as a sum of k i.i.d. exponential random variables, $(X_1/\sum_{i=1}^n X_i, \ldots, X_n/\sum_{i=1}^n X_i)$ and $\sum_{i=1}^n X_i$ are independent. Therefore

$$E\left(\frac{\max\{X_i\}}{\sum_{i=1}^n X_i}\right) = \frac{E(\max\{X_i\})}{E(\sum_{i=1}^n X_i)},$$

and we can relate the maximum to an occupancy distribution, with a known expected value. (See [10].) If this is done we obtain

$$E\left(\frac{\max\{X_i\}}{\sum_{i=1}^n X_i}\right) = \frac{\log n+(k-1)\log\log n+C-\log((k-1)!)+o(n)}{kn}$$

(here $C = 0.57721...$ is Euler's constant), a very accurate approximation for fixed k [7]. If we approximate the gamma distribution by truncation to $[0, M]$ and $M \gg k$, then $\overline{S}$ will be close to this. On the other hand, for any fixed n,

$$\lim_{M \to \infty} E\left(\frac{M}{M + \sum_{i=1}^{n-1} X_i}\right) = 1$$

by the bounded convergence theorem.

Roughly speaking, then, we expect (15) to be good for distributions with significant mass near the upper bound M. Since $\max\{X_i\}$ and $\sum_{i=1}^{n} X_i$ are asymptotically independent [12], there is probably some merit in replacing (15) by

$$\frac{E[\max\{X_i\}]}{M}\hat{S}$$

if this is not the case. We note, however, that this heuristic correction requires knowledge of the expected maximum.

In the next section we will validate the approximation (15) for several distributions and assess its accuracy as a function of n.

3 Validations

We validated approximation (15) for several distributions of X. Clearly, for the trivial case of equal task service times, i.e., $X_i \equiv 1$, $i = 1, \ldots, n$, the approximation is exact since $C_X = 0$, $E[V] = 1$ and $\overline{S} = 1/n$ as required. To validate the approximation for more variable X we used simulation to estimate $\overline{S}$ for uniform, beta, triangular, and truncated exponential distributions for X. Table 1 provides the density, mean, and coefficient of variation for each of these distributions. (Note that Beta(1,1) is equivalent to Uniform(0,1), and Beta(1,2) is a scaled version of a Triangular distribution.) The Truncated Exponential(a, b) distribution is obtained by truncating an exponential with mean a at the point b, and replacing the tail by a mass point. All simulation estimates were for a sample size of 10000 and had 99% confidence intervals with less than 1% half-widths. The sample standard deviation was used to estimate the confidence intervals. All approximate estimates of $\overline{S}$ were computed using (15) after dropping the $O(1/n^2)$ error term.

Figures 1(a) and (b) plot $\overline{S}$ versus n for various settings of the Beta and Truncated Exponential distributions. In Figure 1(a) the curves are in increasing order of C_X. We observe that for the Beta(2,1) and

Table 1: Distributions used in Validations

X	**Density** $f_X(x)$	$E[X]$	C_X^2
Beta(a,b)	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$, $0<x<1$	$\frac{a}{a+b}$	$\frac{b}{a(a+b+1)}$
Triangular(b)	$\frac{2(b-x)}{b^2}$, $0<x<b$	$\frac{b}{3}$	$\frac{1}{2}$
Truncated Exponential(a,b)	$\begin{cases} \frac{1}{a}e^{-x/a}, & 0<x<b, \\ e^{-b/a}\,\delta(b), & x=b \end{cases}$	$a-be^{-b/a}$	$\frac{2\{a-(a+b)e^{-b/a}\}}{(a-be^{-b/a})^2}-1$

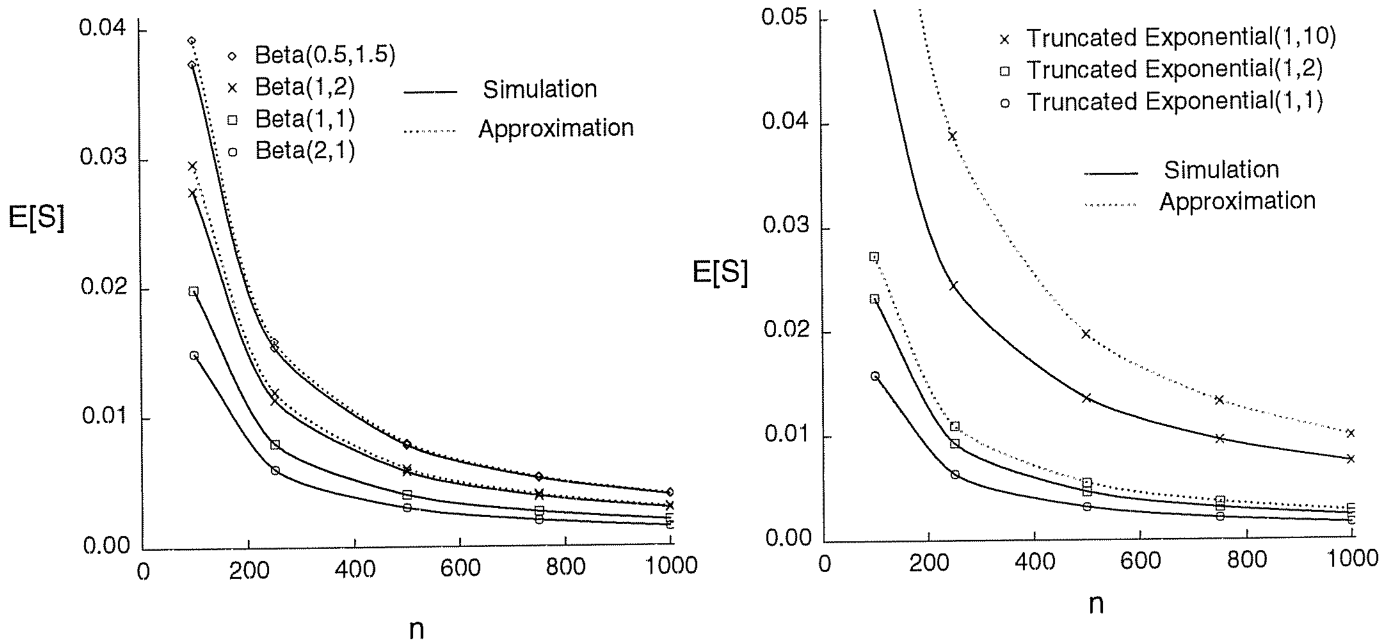
In all cases above a, $b>0$

Beta(1,1) distributions approximation estimates of $\overline{S}$ coincide with the corresponding simulation estimates. As C_X increases the accuracy of the approximation degrades slightly when $n<500$ as seen from the curves for Beta(1,2) and Beta(0.5,1.5). For $n>500$ the approximation estimates are almost identical to the simulation estimates. From Figure 1(b) we observe that for the Truncated Exponential(1,1) distribution the approximation is very accurate. When the truncation point of the exponential is increased from 1 to 2 the approximation is less accurate for $n<1000$ and when b is increased still further to 10 the approximation is rather poor for $n<1000$. For this choice of b we found the approximation to be quite accurate for $n>2500$ (not shown). The reason for the poor accuracy for the Truncated Exponential(1,10) distribution when $n<1000$ is that our approximation is based on the fact that $\max\{X_1,\ldots,X_n\}\to M$ as $n\to\infty$. For the truncated exponential distribution n has to be quite large for $P[\max\{X_1,\ldots,X_n\}>M-\epsilon]>1-\delta$, for small positive ϵ and δ. For example, for $n=1000$, $M=10$, and $\epsilon=0.05$, $P[\max\{X_1,\ldots,X_n\}>9.95]=0.05$. In contrast, for a Beta(1,1) random variable, $n=100$, $M=1$, and $\epsilon=0.05$, $P[\max\{X_1,\ldots,X_n\}>0.95]=0.994$, which is why the approximation is extremely accurate for the Beta(1,1) curve for all $n\geq 100$.

One might conclude from the experiments that $\hat{S}$, for which (15) is an approximation, always overestimates $\overline{S}$. The following example shows that this is not the case. Let $X=1/2$ with probability $1/4$, and 1 with probability $3/4$. For $n=2$ we have $\overline{S}=0.5625$, whereas $\hat{S}=0.54166\ldots$

Appendix

In this appendix we collect some technical results from probability theory.



(a) $X \sim \text{Beta}(a, b)$

(b) $X \sim \text{Truncated Exponential}(a, b)$

Figure 1: Validations: $\overline{S}$ versus n

1. Sampling Theory. For any random variable ξ, let

$$\mu_k(\xi) := E\left[(\xi - E[\xi])^k\right]$$

be its k-th central moment. For sample means, we have

$$\mu_k\left(\frac{\sum_{i=1}^m V_i}{m}\right) = E\left[\left(\frac{\sum_{i=1}^m (V_i - E[V])}{m}\right)^k\right].$$

Let $\mu_k := \mu_k(V)$ denote the k-th central moment of the parent distribution. Cramèr [3, page 345] gives the following formulas for the i.i.d. case:

$$\begin{aligned}
\mu_1\left(\frac{\sum_{i=1}^m V_i}{m}\right) &= 0; \\
\mu_2\left(\frac{\sum_{i=1}^m V_i}{m}\right) &= \frac{\mu_2}{m}; \\
\mu_3\left(\frac{\sum_{i=1}^m V_i}{m}\right) &= \frac{\mu_3}{m^2};
\end{aligned}$$

$$\mu_4\left(\frac{\sum_{i=1}^m V_i}{m}\right) = \frac{3\mu_2^2}{m^2} + \frac{\mu_4 - 3\mu_2^2}{m^3},$$

and observes that for fixed k, $\mu_k(\sum_{i=1}^m V_i/m) = O(m^{-\lceil k/2 \rceil})$. Higher moments could be computed using a procedure outlined by Kendall and Stuart [8], although the formulas become increasingly complex as k grows. In particular, it is not obvious how the constant in the above O-result depends on k.

2. Large Deviations. For a random variable ξ, let $\Psi_\xi(t) := E[e^{t\xi}]$ be its moment generating function. The following result appears in Chernoff [4]: If $V_1, V_2, \ldots, V_m$ are i.i.d. with the same distribution as ξ and $a < E[\xi]$, then

$$\Pr[\sum_{i=1}^m V_i \leq am] \leq \left(\inf_{\lambda \geq 0}\left\{e^{a\lambda}\Psi_\xi(-\lambda)\right\}\right)^m.$$

We must check that the infimum is actually less than 1. To do this, note that $e^{a\lambda}\Psi_\xi(-\lambda)$ is the moment generating function of $a - \xi$. This is 1 for $\lambda = 0$, and its derivative at 0 is $E[a - \xi] < 0$. Since moment generating functions are analytic, there will be some $\lambda > 0$ where the function has a value less than 1.

3. Convergence of Moments. Let $V_1, \ldots, V_m$ be i.i.d. with $0 \leq V_i \leq 1$, and consider the normalized sum

$$Y_m := \frac{\sum_{i=1}^m (V_i - E[V])}{\sigma_V \sqrt{m}}.$$

As is well-known, $Y_m \to N(0,1)$ in distribution, but we need the corresponding relation for moments. We will use the *Second Limit Theorem* from Kendall and Stuart [8, page 115], which states that for a sequence of distribution functions $\{F_m(x)\}$ that converge to the distribution function $G(x)$, the k^{th} central moment of $F_m(x)$, $\mu_k(m)$, converges to the k^{th} central moment, λ_k, of $G(x)$ (assuming that $\mu_k(m)$ exists for all $m > m_0$ and for all $k \geq 0$) provided that $\mu_k(m)$ is bounded above by some constant A_k that is independent of m.

Clearly $E[Y_m^k]$ exists for $m \geq 1$ since the V_i are bounded. We need to show that $E[Y_m^k] \leq A_k$ for some A_k independent of m. To do this we use the following result by Whittle [14]:

Proposition. *Let $W_m = \sum_{i=1}^m b_i U_i$, where $U_1, \ldots, U_m$ are i.i.d. with mean zero. Let $\gamma_i(k) = E[|U_i|^k]^{1/k}$. Then for $k \geq 2$,*

$$E[|W_m|^k] \leq 2^k C(k)\left(\sum_{i=1}^m b_i^2 \gamma_i^2(k)\right)^{k/2},$$

where

$$C(k) := \frac{2^{k/2}}{\sqrt{\pi}}\,\Gamma((k+1)/2).$$

To apply Whittle's result we set $U_i = (V_i - E[V])/(\sigma_V\sqrt{m})$ and $b_i = 1$. Thus $Y_m = \sum_{i=1}^m U_i$, and

$$\gamma_i(k) = E[|U_i|^k]^{1/k} = E\left[\left|\frac{V_i - E[V]}{\sigma_V\sqrt{m}}\right|^k\right]^{1/k} = \frac{1}{\sigma_V\sqrt{m}}\,E\left[|V_i - E[V]|^k\right]^{1/k},$$

which is independent of i as $V_1, \ldots, V_m$ are i.i.d. As a result,

$$\begin{aligned}
E[|Y_m|^k] &\leq 2^k C(k)\left(\sum_{i=1}^m \gamma_i^2(k)\right)^{k/2} \\
&= 2^k C(k)\left(\frac{1}{\sigma_V^2\,m}\sum_{i=1}^m E\left[|V_i - E[V]|^k\right]^{2/k}\right)^{k/2} \\
&= 2^k C(k)\left(\frac{1}{\sigma_V^2\,m}\,m\,E\left[|V_i - E[V]|^k\right]^{2/k}\right)^{k/2} \\
&= 2^k C(k)\,\frac{1}{\sigma_V^k}\,E\left[|V_i - E[V]|^k\right],
\end{aligned}$$

which is independent of m. Therefore, by the second limit theorem,

$$E[Y_m^k] \rightarrow k^{th} \text{ moment of a } N(0,1) = \begin{cases} 0, & k \text{ odd;} \\ \frac{k!}{2^{k/2}(k/2)!}, & k \text{ even.} \end{cases}$$

whenever $k \geq 1$. Whittle's result also gives an explicit bound for the k-th absolute moment of Y_m; we have

$$E[|Y_m|^k] \leq \frac{2^{3k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}\sigma_V^k}\,E\left[|V - E[V]|^k\right].$$

∎

References

[1] D. Z. Arov and A. A. Bobrov. The Extreme Terms of a Sample and Their Role in the Sum of Independent Variables. *Teor. Verojatnost. i Primen.* 5 (1960), 377-396. [In Russian; English translation in *Theory of Probability and Applications.*]

[2] T. L. Chow and J. L. Tuegels, The sum and maximum of i.i.d. random variables, Proc. *Second Prague Symposium on Asymptotic Statistics*, (1979), 81-92.

[3] H. Cramér, *Mathematical Methods of Statistics*. Princeton Univ. Press, Princeton 1946.

[4] H. Chernoff. A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations. *Annals of Math. Stat.*, 23 (1952), 493-507.

[5] D. A. Darling. The Influence of the Maximum Term in the Addition of Independent Random Variables. *Trans. Amer. Math. Soc.* 73 (1952), 95-107.

[6] D. A. Darling. On a Test for Homogeneity and Extreme Values. *Ann. Math. Stat.* 23 (1952), 450-456.

[7] P. Erdős and A. Rényi, On a Classical Problem of Probability Theory. *MTA Mat. Kut. Inst. Kozl.* 6A (1961), 215-220.

[8] M. Kendall and A. Stuart. *The Advanced Theory of Statistics. Volume 1: Distribution Theory.* Second Edition, Charles Griffin & Co. Ltd., London 1963.

[9] S. Leutenegger. Issues in multiprogrammed multiprocessor scheduling. Ph.D. Thesis, Technical Report #954, Computer Sciences Dept., Univ. of Wisconsin-Madison, August 1990.

[10] G. Lewandowski, A. Condon, and E. Bach. Realistic analysis of parallel dynamic programming algorithms. Technical Report #1116, Computer Sciences Dept., Univ. of Wisconsin-Madison, October 1992. (Abstract in Proc. 1993 *SIGMETRICS*, 268-269.)

[11] S. Majumdar, D. Eager, and R. Bunt. Scheduling in multiprogrammed parallel systems. *ACM SIGMETRICS Conf. and Performance Evaluation Review 16*, 1 (May 1988), 104-113.

[12] A. Rosengard. Étude des Lois-limites Jointes et Marginales de la Moyenne et des Valeurs Extrêmes d'un Énchantillon. *Publ. Inst. Statist. Univ. Paris*, 11 (1963), 3-53.

[13] J. Strikwerda and R. Mansharamani. Evaluating the Mean Completion Time of a Fork-Join Barrier Synchronization. Technical Report #1153, Computer Sciences Department, University of Wisconsin-Madison, May 1993. *Submitted for publication.*

[14] P. Whittle. Bounds for the Moments of Linear and Quadratic Forms in Independent Variables. *Theory of Probability and Its Applications V*, 3 (1960), 302-305.