WORKING PAPER

ALFRED P. SLOAN SCHOOL OF MANAGEMENT

ON THE RATE OF CONVERGENCE TO OPTIMALITY
OF THE LPT RULE

by

J.B.G. Frenk[*]

A.H.G. Rinnooy Kan[**]

Working Paper # 1610-84

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

ON THE RATE OF CONVERGENCE TO OPTIMALITY
OF THE LPT RULE

by

J.B.G. Frenk[*]

A.H.G. Rinnooy Kan[**]

Working Paper # 1610-84

Abstract

The LPT rule is a heuristic method to distribute jobs among identical machines
so as to minimize the makespan of the resulting schedule. If the processing
times of the jobs are assumed to be independent identically distributed random
variables, then (under a mild condition on the distribution) the absolute
error of this heuristic is known to converge to 0 almost surely. In this
note we analyse the asymptotic behaviour of the absolute error and its first
and higher moments to show that under quite general assumptions the speed of
convergence is proportional to appropriate powers of (log log n)/n and 1/n.
Thus, we strengthen and extend earlier results obtained for the uniform
and exponential distribution.

(*)   Department of Industrial Engineering and Operations Research, University
       of California, Berkeley

(**)   Econometric Institute, Erasmus University, Rotterdam/Sloan School of
       Management, M.I.T., Cambridge

ACKNOWLEDGMENTS

1. INTRODUCTION

Suppose that n jobs with processing times $p_1, \ldots, p_n$ have to be distributed among m uniform machines. Let s_i be the speed of machine i (i = 1, ..., m). If the sum of the processing times assigned to machine i is denoted by $Z_n(i)$ (i = 1, ..., m), then a common objective is to minimize the makespan $Z_n^{(m)} = \max_i \{Z_n(i)/s_i\}$. For this NP-hard problem many heuristics have been proposed and analyzed; we refer to [Graham et al. 1979; Rinnooy Kan 1984] for a survey. Among them, the LPT rule in which jobs are assigned to the first available machine in order of decreasing p_j is a particularly simple and attractive one. The value $Z_n^{(m)}(LPT)$ produced by this rule is related to the optimal solution value $Z_n^{(m)}(OPT)$ for the case that $s_i = 1$ for all i by [Graham 1969]

$$\frac{Z_n^{(m)}(LPT)}{Z_n^{(m)}(OPT)} \leq \frac{4}{3} - \frac{1}{3m} \quad . \tag{1}$$

Computational evidence, however, suggests that this worst case analysis is unnecessarily pessimistic in that problem instances for which (1) is satisfied as an equality appear to occur only rarely.

To achieve a better understanding of this phenomenon, let us assume the processing times p_j (j = 1, ..., n) to be independent, identically distributed random variables. The relation between the random variables $\underline{Z}_n^{(m)}(OPT)$ and $\underline{Z}_n^{(m)}(LPT)$ can then be subjected to a probabilistic analysis. In [Frenk & Rinnooy Kan 1984] it was shown that (under mild conditions on the distribution of the p_j) the absolute error

$$\underline{Z}_n^{(m)}(LPT) - \underline{Z}_n^{(m)}(OPT) \tag{2}$$

converges to 0 almost surely as well as in expectation. Thus, the heuristic is asymptotically optimal in a strong (absolute rather than relative) sense, which provides an explanation for its excellent computational behaviour.

In [Frenk & Rinnooy Kan 1984], the speed at which the absolute error converges
to 0 was analyzed for the special cases of the uniform and exponential
distribution respectively. Here we extend and generalize the results for
almost sure convergence and convergence in expectation by showing that for
a large class of distributions (essentially those with $F(x) = x^a$ $(0 \_ x \leq 1$,
$0 \quad a \quad \cdot))$, this speed is proportional to an appropriate power of $(\log \log n)/n$
and $1/n$ respectively. This implies that, although the
optimality of the LPT rule could only be established asymptotically,
the convergence of the absolute error to 0 at least occurs reasonably fast.
In some sense, to be explained later, these results are the best possible
ones obtainable for this heuristic.

The main result for the case of almost sure convergence, is described and
proved in Section 2. The case of convergence in expectation is dealt with
in Section 3, where we bound first and higher moments of the expected absolute
error. Some extensions and conjectures are briefly examined in Section 4.

2. ALMOST SURE CONVERGENCE

In [Frenk & Rinnooy Kan 1984], it is shown that the absolute error of the
LPT rule (2) is bounded (up to a multiplicative constant) by

$$D_n(\alpha) = \max_{1 \leq k \leq n} \left\{ \underline{p}_{k:n} - \frac{1}{\alpha} \sum_{j=1}^{k} \underline{p}_{j:n} \right\} \tag{3}$$

where $\underline{p}_{1:n} \leq \underline{p}_{2:n} \leq \dots \leq \underline{p}_{n:n}$ are the <u>order statistics</u> of the processing
times and $\alpha = 1 + (m - 1)s_1/s_m$. Let us assume that the distribution function
of the processing times is given by

$$F(x) = x^a \quad (0 \leq x \leq 1, \ 0 < a < \infty) \tag{4}$$

In that case, $\underline{p}_{k:n} \overset{d}{=} \underline{U}_{k:n}^b$, where $\underline{U}_{k:n}$ $(k = 1, \dots, n)$ are the order
statistics of n independent random variables uniformly distributed on $[0, 1]$,
and $b = 1/a$.

-2-

To study the asymptotic behaviour of $\underline{D}_n(\alpha)$, let us define the random variable $\underline{T}_n$ to be the index $p \in \{1, \ldots, n\}$ for which the maximum in (3) is actually achieved. Hence, $\underline{T}_n = p$ implies that

$$\underline{U}^b_{p:n} - \frac{1}{\alpha} \Sigma^p_{j=1} U^b_{j:n} \geq \underline{U}^b_{k:n} - \frac{1}{\alpha} \Sigma^k_{j=1} U^b_{j:n} \qquad (k = 1, \ldots, p-1) \quad (5)$$

i.e., that

$$\alpha \, \underline{U}^b_{k:n} + \Sigma^{p-1}_{j=k+1} \underline{U}^b_{j:n} - (\alpha - 1) \underline{U}^b_{p:n} \leq 0 \qquad (k = 1, \ldots, p-1) \quad (6)$$

so that (by addition of these inequalities)

$$\Sigma^{p-1}_{k=1} (\alpha + k - 1) \underline{U}^b_{k:n} \leq (\alpha - 1)(p - 1) \underline{U}^b_{p:n} \quad (7)$$

Thus, $\Pr\{\underline{T}_n = p\}$ is bounded (from above) by the probability of (7).

Now, it is easily verified that

$$\Pr\{\underline{U}^b_{k:n} \leq x_k (k = 1, \ldots, p - 1) \mid \underline{U}^b_{p:n} = y\} =$$
$$\Pr\{(\underline{U}_{k:p-1}(y^a))^b \leq x_k \ (k = 1, \ldots, p - 1)\} \quad (8)$$

where, for any $z \in [0, 1]$, $U_{k:n}(z)$ $(k = 1, \ldots, p - 1)$ are the order statistics of $p - 1$ independent random variables, uniformly distributed on $[0, z]$ [Karlin & Taylor 1981, p. 103].

Let $F_p(y) = \Pr\{\underline{U}^b_{p:n} \leq y\}$. Then (7) and (8) imply that

$$\Pr\{\underline{T}_n = p\} \leq$$
$$\int^1_0 \Pr\{\Sigma^{p-1}_{k=1}(\alpha + k - 1)(\underline{U}_{k:p-1}(y^a))^b \leq (\alpha - 1)(p - 1)y\} \, F_p(dy) \quad (9)$$

Since $(\underline{U}_{k:p-1}(y^a))^b / y \overset{d}{=} \underline{U}^b_{k:p-1}$, (9) is bounded by

$$\Pr\{\Sigma^{p-1}_{k=1} (\alpha + k - 1) \underline{U}^b_{k:p-1} \leq (\alpha - 1)(p - 1) \} \quad (10)$$

Now Lemma 1 in the Appendix implies that for certain constants $C = c(\alpha)$, $c = c(\alpha)$

$$\Pr\{\underline{T}_n = p\} \leq C \, e^{-cp} \quad (11)$$

To derive our main result from the Bovel-Cantelli lemma, we now use (11) to bound

$$\Pr\{\underline{D}_n(\alpha) \geq \left[\frac{D \, \log_2 n}{n} \right] \} \quad (12)$$

-3-

(where D is a constant to be chosen later and $\log_2 n = \log \log n$) by

$$\Pr\{T_n \geq \log n\} +$$

$$+ \Pr\{\max_{1 \leq k \leq \log n} \{U_{k:n}^b - \frac{1}{\alpha} \Sigma_{j=1}^k U_{j:n}^b\} \geq (\frac{D \log_2 n}{n})^b \} \qquad (13)$$

The first term in (13) is $O(n^{-c})$ from (11). We again condition on the value $U_{\log n:n}^b$ corresponding to the largest order statistic being greater or smaller than $(2(\log n)/n)^b$, to bound the second term by

$$\Pr\{U_{\log n:n}^b \geq (\frac{2 \log n}{n})^b\} +$$

$$+ \int_0^{(2 \log n/n)^b} \Pr\{\max_{1 \leq k \leq \log n} U_{k:n}^b - \frac{1}{\alpha} \Sigma_{j=1}^k U_{j:n}^b \geq (\frac{D \log_2 n}{n})^b |$$

$$| U_{\log n:n}^b = y\} \; F_{\log n}(dy) \qquad (14)$$

The first term in (14) is $O(n^{-\frac{1}{2}})$ (cr [de Haan & Taconis 1979]). To bound the second term, we observe that the term within the integral is bounded for every $y \in (0, 1)$ by (cf (8))

$$\Pr\{\max_{1 \leq k \leq \log n - 1} U_{k:\log n - 1}^b - \frac{1}{\alpha} \Sigma_{j=1}^k U_{j:\log n - 1}^b \geq \frac{1}{y}(\frac{D \log_2 n}{n})^b\}$$

$$+ \Pr\{(1 - \frac{1}{\alpha}) - \frac{1}{\alpha} \Sigma_{j=1}^{\log n - 1} U_j^b \geq \frac{1}{y}(\frac{D \log_2 n}{n})^b\} \qquad (15)$$

so that the integral itself is bounded by

$$\Pr\{\max_{1 \leq k \leq \log n - 1} \{U_{k:\log n - 1}^b - \frac{1}{\alpha} \Sigma_{j=1}^k U_{j:\log n - 1}^b\} \geq \left(\frac{D \log_2 n}{2 \log n}\right)^b +$$

$$+ \Pr\{\frac{1}{\alpha} \Sigma_{j=1}^{\log n - 1} U_j^b < 1 - \frac{1}{\alpha}\} \qquad (16)$$

The second probability in (16) converges exponentially to 0 (in log n). We again bound the first probability by conditioning on the index $T(\log n - 1)$ (where the maximum is attained) being greater or smaller than $d \log_2 n$, for a constant d still to be chosen. From (11), the probability of the former event is $O((\log n)^{-cd})$. The remaining conditional probability is bounded by

-4-

$$\Pr\{\underline{U}^b_{-d \ \log_2 n : \log n - 1} \geq (\frac{D \ \log_2 n}{2 \ \log \ n})^b\} \tag{17}$$

For $d = D/4$, this term is $0((\log \ n)^{-D/16})$ (cf [de Haan & Taconis 1979]).

Collecting all our upper bounds on (12), we conclude that, if

$D = 2 \ \max \ \{16, \ 4/c\}$

$$\Pr\{\underline{D}_n(\alpha) \geq (\frac{D \ \log_2 n}{n})^b\} = 0(1/(\log n)^2) \tag{18}$$

Define $k_n = e^n$. The Borel-Cantelli lemma implies immediately that

$$\lim \ \sup_{n \to \infty} (\frac{k_n}{\log_2 \ k_n})^b \ \underline{D}_{k_n}(\alpha) < \infty \qquad (a.s.) \tag{19}$$

We show in the Appendix (Lemma 2) that $\underline{D}_n(\alpha)$ is almost surely nonincreasing

in n. It follows that

$$\lim \ \sup_{n \to \infty} (\frac{n}{\log_2 n})^b \ \underline{D}_n(\alpha) < \infty \qquad (a.s.) \tag{20}$$

and we have proved the main result of this section.

Theorem 1. If the distribution function of the processing times equals

$F(x) = x^a$ $(0 \leq x \leq 1, \ 0 < a < \infty)$, then

$$\lim \ \sup_{n \to \infty} (\frac{n}{\log_2 n})^{1/a} \ (\underline{Z}^{(m)}(LPT) \ - \ Z_n^{(m)}(OPT)) < \infty \qquad (a.s.)$$

This speed of convergence result is the best possible one that can be

derived from the upper bound (3), as can be seen from the fact that

$$\Pr\{\underline{U}_{1:n} \geq \frac{\log_2 n}{n} \qquad i.o.\} = 1 \tag{21}$$

It can be shown [Karp 1983] that the speed of convergence to optimality for

the LPT rule is at least $1/n$ for the case that $a = 1$. In the next section,

we shall see that this lower bound is also an upper bound when we consider

convergence in expectation.

3. CONVERGENCE IN EXPECTATION

Again, we assume that $F(x) = x^a$ ($0 \leq x \leq 1$, $0 < a < \infty$). With T_{-n} as defined before,

$$E(D_{-n}(x)^q) \leq \Pr\{T_{-n} = n\} +$$
$$+ E(\max_{1 \leq k \leq n - 1} \{U^b_{-k:n} - \frac{1}{\alpha} \Sigma^k_{j=1} \ U^b_{-j:n}\})^q \tag{22}$$

As before we condition on the value of the largest order statistic to bound the second term by

$$E(E(\max_{1 \leq k \leq n - 1} \{U^b_{-k:n} - \frac{1}{\alpha} \ \Sigma^k_{j=1} \ U^b_{-j:n}\})^q \mid U_{-n:n}) =$$

$$E(U^{qb}_{-n:n} \ E(\max_{1 \leq k \leq n - 1} \{ \frac{U_{-k:n}}{U_{-n:n}}^b - \frac{1}{\alpha} \Sigma^k_{j=1} \ \frac{U_{-j:n}}{U_{-n:n}}^b \})^q \mid U_{-n:n}) =$$

$$E(U^{qb}_{-n:n}) \ E(\max_{1 \leq k \leq n - 1} \{U^b_{k:n - 1} - \frac{1}{\alpha} \Sigma^k_{j=1} \ U^b_{j:n-1}\})^q \tag{23}$$

Hence, for n sufficiently large, (11) and (23) together imply that

$$E(D_{-n}(x)^q) \leq e^{-cn} + \frac{n}{n + qb} E(D_{-n - 1}(\alpha)^q) \tag{24}$$

Let $h_n = (n + 1)^{qb} E(D_{-n}(\alpha)^q)$. Then (24) implies that

$$h_n \leq (n + 1)^{qb} \ e^{-cn} + e^{(qb/n)^2} \ h_{n-1} \tag{25}$$

This implies that h_n is bounded by a constant and we have proved the main result of this section.

Theorem 2. If the distribution function of the processing times $F(x) = x^a$ ($0 \leq x \leq 1$, $0 < a < \infty$), then

$$\lim \sup_{n \to \infty} (n^{q/a} E((Z_n^{(m)}(LPT) - Z_n^{(m)}(OPT))^q) \times \infty$$

Identical results for the case that $s_i = 1$ for all i are derived in a different fashion in [Boxma 1984]. Again, they are the sharpest possible ones in the sense of the previous section. It is worth noting that this is the first time that bounds on higher moments have been derived for a heuristic of this nature.

4. EXTENSIONS AND CONCLUDING REMARKS

Theorems 1 and 2 can both be extended to the case that $F(x) = \Theta(x^a)$, i.e., there exist positive constants ε, L and U such that

$$Lx^a \leq F(x) \leq Ux^a \tag{26}$$

for $x \in [0, \varepsilon)$.

In the case of almost sure convergence, this is done by showing that one may restrict oneself in the maximization (3) to $k \in \{1, \ldots, [\varepsilon n]\}$ (as in [Frenk & Rinnooy Kan 1984]). This maximization involves only the smaller order statistics and for those we are essentially in the situation analyzed in Section 2.

In the case of convergence in expectation, our technique requires that

$$E \underline{p}^{q(1+b)+1} < \infty \tag{27}$$

for the extension of Theorem 3 to hold. We strongly suspect, however, that this condition is not essential. Details of the proofs for these results are available from the authors.

These unusually strong results, as well as other recent ones in this area ([Boxma 1984]), confirm the remarkable amenability of the LPT rule to a probabilistic analysis. Extensions to other priority rules involving order statistics of processing times seem feasible and interesting.

APPENDIX

Lemma 1. For every $\beta \geq 1$ there are positive constants $C = C(\beta)$

and $c = c(\beta)$ such that

$$\Pr \{\Sigma_{k=1}^{m} (\beta + k) \underset{-k:m}{U}^{1/a} \leq \beta m\} \overset{\leq}{=} C \cdot e^{-cm}$$

Proof. If $a \geq 1$, then $\underset{-k:m}{U}^{1/a} \geq \underset{-k:m}{U}$ a.s.

Also, if $a < 1$, we obtain from Hölder's inequality [Goffman & Pedrick

1965, p. 2] that (take $p = 1/a$, $q = 1/(1-a)$, $y_k = (\beta + k)^a \underset{-k:m}{U}$, $x_k = 1$)

$$\Sigma_{k=1}^{m} (\beta + k)^a \underset{-k:m}{U} \underset{-}{\leq} m^{1-a} (\Sigma_{k=1}^{m} (\beta + k) \underset{-k:m}{U}^{1/a})^a \qquad \text{a.s.} \qquad (A.1)$$

This implies that

$$\Pr \{\Sigma_{k=1}^{m} (\beta + k) \underset{-k:m}{U}^{1/a} \underset{-}{\leq} \beta m\}$$

$$= \Pr \{(\Sigma_{k=1}^{m} (\beta + k) \underset{-k:m}{U}^{1/a})^a \underset{-}{\leq} \beta^a m^a\}$$

$$\underset{-}{\leq} \Pr \{\Sigma_{k=1}^{m} (\beta + k)^a \underset{-k:m}{U} \underset{-}{\leq} \beta^a m\}. \qquad (A.2)$$

Hence we consider the distribution of

$$\Sigma_{k=1}^{m} (\beta + k)^a \underset{-k:m}{U} \quad (a \leq 1)$$

Since $(\underset{-1:m}{U}, \ldots, \underset{-m,m}{U})^d = (\underset{-1}{S}/\underset{-m+1}{S}, \ldots, \underset{-m}{S}/\underset{-m+1}{S})$ with

$\underset{-i}{S} = \Sigma_{j=1}^{i} \underset{-j}{V}$ and $\underset{-j}{V}$ independent exponentially distributed random

variables with parameter $\lambda = 1$ ($j = 1, \ldots, m$) ([Karlin & Taylor 1981,

p. 103]), we can rewrite the right hand side of (A.2) as

$$\Pr\{\Sigma_{k=1}^{m+1} (\beta + k)(\underset{-m+1}{S} - \underset{-k}{S}) \geq \underset{-m+1}{S} (\Sigma_{k=1}^{m+1} (\beta + k)^a - \beta^a m)\}$$

Now for every $\varepsilon > 0$, there exists some $m_0 = m_0(\varepsilon)$ such that, for every $m \geq m_0(\varepsilon)$, the above probability is bounded from above by

$$\Pr \{\Sigma_{\ell=1}^{m+1} (1 + \ell + \beta)^{a+1} \underline{V}_\ell \geq ((1 - \varepsilon)m)^{a+1} \underline{S}_{m+1}\}$$

$$= \Pr\{\Sigma_{\ell=1}^{m+1} c_{\ell,m+1} \underline{V}_\ell \geq 0\}$$

with

$$c_{\ell,m+1} = ((1 + \ell + \beta)/(1 - \varepsilon)m)^{a+1} - 1 \tag{A.4}$$

Clearly $-1 < c_{\ell,m+1} \leq ((m + 2 + \beta)/(1 - \varepsilon)m)^{a+1} - 1$ $(\ell = 1,\ldots, m + 1)$

and this implies for $\lambda \varepsilon [0, 1/2 \cdot (2(1 - \varepsilon)/3)^{a+1}]$

and $m \geq \max(m_0(\varepsilon), 2(\beta + 2))$ that

$$\Pr \{\Sigma_{\ell=1}^{m+1} c_{\ell,m+1} \underline{V}_\ell \geq 0\} \leq E(\exp(\lambda \Sigma_{\ell=1}^{m+1} c_{\ell,m+1} \underline{V}_\ell))$$

$$= \pi_{\ell=1}^{m+1} 1/(1 - \lambda c_{\ell,m+1}) . \tag{A.5}$$

From the Taylor expansion of $\log (1 + x)$ around $x = 0$, we then show that the above term is bounded by

$$\exp(\lambda \Sigma_{\ell=1}^{m+1} c_{\ell,m+1} + \lambda^2 \Sigma_{\ell=1}^{m+1} c_{\ell,m+1}^2) .$$

Since

$$\lim_{m \to \infty} (\Sigma_{\ell=1}^{m+1} c_{\ell,m+1})/m = 1/((a + 1)(1 - \varepsilon)^{a+1}) \tag{A.6}$$

and

$$\lim_{m \to \infty} (\Sigma_{\ell=1}^{m+1} c_{\ell,m+1}^2)/m = 1/((2a + 3)(1 - \varepsilon)^{2a+2})$$

$$- 2/((a + 2)(1 - \varepsilon)^{a+1}) + 1, \tag{A.7}$$

the desired result follows from the appropriate choice of positive values for λ and ε.

Lemma 2

$$\sum_{n=1}^{\infty} \quad Pr \; \{\underline{D}_{n+1} \; (\;) > \underline{D}_n \; (\;)\} < \infty.$$

Proof. It is easy to verify that $D_{n+1}(\cdot)\leq D_n(\cdot)$ unless (perhaps) if the new processing time is larger than all the previous ones.

Hence,

$$Pr \; \{\underline{D}_{n+1}(\cdot) > \underline{D}_n(\cdot)\} \leq Pr \; \{p_{n+1} - \frac{1}{\alpha} \sum_{j=1}^{n+1} p_j \cdot 0\}$$

$$= \int_0^{\infty} \; F^{n*}((\alpha-1)y) \; F(dy) \tag{A.8}$$

so that

$$\Sigma_{n=1}^{\infty} \quad Pr \; \{\underline{D}_{n+1}(\cdot) > \underline{D}_n(\cdot)\} \leq \int_0^{\infty} \; U((\alpha-1)y)F(dv) \tag{A.9}$$

with $U(x)=\Sigma_{n=1}^{\infty} \; F^{n*}(x)$ the renewal function ([Feller 1971; Van Dulst & Frenk 1984]). The result now follows from

$$\lim_{x\to\infty} \quad \frac{U(x)}{x} = \int_0^{\infty} \; x \; F(dx) < \infty \tag{A.10}$$

and the local boundedness of $U(x)$ on $(0, \infty)$.

REFERENCES

Boxma, O.J., 'A Probabilistic Analysis of the LPT Scheduling Rule', Technical report, Department of Mathematics, University of Utrecht.

de Haan,L., Taconis-Haantjes, E., 'On Bahadur's Representation of Sample Quartiles', Ann. Inst. Statist. Math. (vol 31), 1979, 299-307.

Feller, W., An introduction to probability theory and its applications, Vol. 2., John Wiley, New York, 1971.

Frenk, J.B.G., Rinnooy Kan, A.H.G., 'The asymptotic optimality of the L.P.T. rule', Technical report, Econometric Institute, Erasmus University, Rotterdam.

Goffman, C., Pedrick, G., First course in functional analysis, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1965.

Karlin, S., Taylor, H.M., A second course in stochastic processes, Academic Press, New York, 1981.

Karp, R.M., Private communication, 1983.

Rinnooy Kan, A.H.G., 'Approximation algorithms - an introduction', Technical report, Erasmus University, Rotterdam, 1984.

Van Dulst, D., Frenk, J.B.G., 'On Banach algebras, subexponential distributions and renewal theory', Technical report 84-20, Mathematical Institute, University of Amsterdam.

BASEMENT

Date Due