# Heuristics for Multimachine Scheduling Problems with Earliness and Tardiness Costs

A. Federgruen • G. Mosheiov
*Graduate School of Business, Columbia University, New York, New York 10027*
*School of Business Administration and Department of Statistics, Hebrew University, Jerusalem*

We consider multimachine scheduling problems with earliness and tardiness costs. We first analyze problems in which the cost of a job is given by a general nondecreasing, convex function $F$ of the absolute deviation of its completion time from a (common) unrestrictive due-date, and the objective is to minimize the sum of the costs incurred for all $N$ jobs. (A special case to which considerable attention is given is the completion time variance problem.)

We derive an easily computable lower bound for the minimum cost value and a simple "Alternating Schedule" heuristic, both of which are computable in $O(N \log N)$ time. Under mild technical conditions with respect to $F$, we show that the worst case optimality (accuracy) gap of the heuristic (lower bound) is bounded by a constant as well as by a simple function of a single measure of the dispersion among the processing times. We also show that the heuristic (bound) is asymptotically optimal (accurate) and characterize the convergence rate as $O(N^{-2})$ under very general conditions with respect to the function $F$. In addition, we report on a numerical study showing that the average gap is less than 1% even for problems with 30 jobs, and that it falls below 0.1% for problems with 90 or more jobs. This study also establishes that the empirical gap is almost perfectly proportional with $N^{-2}$, as verified by a regression analysis.

Finally, we generalize the heuristic to settings with a possibly restrictive due date and general asymmetric, and possibly nonconvex, earliness and tardiness cost functions and demonstrate its excellent performance via a second numerical study.
(*Multimachine Scheduling; Earliness and Tardiness Costs; Heuristics; Worst Case; Asymptotic and Probabilistic Analysis*)

Thanks to the widespread interest in Just-in-Time production strategies (e.g., Hall 1983, Monden 1983, Blackburn 1991) much attention has been given to the problem of scheduling jobs with the objective of completing them as close as possible to a common due date. Almost all the literature on this type of scheduling problems with earliness and tardiness costs confines itself to *single* machine settings. We refer to Baker and Scudder (1990) for a survey of the single machine literature up to 1990.

This paper addresses Just-in-Time scheduling problems on multiple parallel machines. With parallel machines, only the problem of minimizing a linear function of the deviations of the jobs' completion times with respect to the due date has been satisfactorily solved. Sundaraghavan and Ahmed (1984) and Hall (1986) provided an $O(N \log N)$ algorithm, with $N$ the number of jobs, to minimize the *sum* of absolute deviations provided the due date is *nonrestrictive*, i.e., it is sufficiently large to have no impact on the optimal assignment of jobs to machines and their sequences on these machines. Emmons (1987) extended this result to the case where the deviations of the completion times of the *tardy* jobs and those of the *early* jobs are weighted with two different factors and to settings with uniform machines, i.e., machines with different speeds. Finally, Kubiak et al. (1990) showed that for unrelated machines, any lin-

0025-1909/96/4211/1544$01.25

ear function of the jobs' deviations can be optimized (in polynomial time) by solving a transportation problem. (With unrelated machines, each job's processing time depends in a general way on its identity and the machine to which it is assigned.)

The restriction to *linear* cost structures and nonrestrictive due dates is not surprising: when the due date is restrictive, the problem is NP-complete even in the *single* machine case, and when minimizing the sum of absolute deviations, arguably the simplest of all cost structures, see Hall et al. (1991). No polynomial time algorithms are known for any *nonlinear* cost structures, even with a *single* machine and even when the due date is *nonrestrictive*. This applies even in the special case where the sum of *squared* deviations is minimized or equivalently (see Bagchi et al. 1987) the *variance* of the completion times. (The latter objective has been studied extensively, but only in the single machine case; see §3. Kubiak 1993b showed that the problem is binary NP-hard.) In the single machine case, one has at least a number of *pseudo-polynomial* methods which can be used for general nonlinear cost structures; see Kahlbacher (1992) and Federgruen and Mosheiov (1993).

In this paper, we first address problems in which the cost of each job is given by a common but *general* nondecreasing and convex function $F$ of the absolute deviation of its completion time from an unrestrictive due date. The objective is to minimize the *sum* of the costs incurred for all jobs. (Special cases include the problem of minimizing the sum of squared deviations and the completion time variance problem.) In §1, we derive a lower bound for the optimum-cost value, and a simple "Alternating Schedule" heuristic, both of which can be computed in $O(N \log N)$ time. We show that for certain problem instances the lower bound is tight and that the heuristic generates an optimal solution.

For general problem instances, we characterize in §2 the worst case optimality (and accuracy) gap as a simple function of a *single* measure of the dispersion among the processing times. Under mild conditions with respect to the function $F$, the optimality gap is bounded by a constant, independent of any of the model parameters. The bound is asymptotically accurate, and the heuristic is asymptotically optimal, as $N \to \infty$. We also characterize the rate at which the optimality/accuracy gap decreases to zero as a function of $N$ and other model parameters.

In §3 we address the important special case where the completion time variance is to be minimized, and we report on a numerical study showing that the average gap between the lower bound and the heuristic is less than 1% even for problems with $N = 30$ jobs, and less than 0.1% when $N \geq 90$. The above results, important in their own right, directly suggest an effective heuristic for more general cost structures, in particular restrictive due dates and *asymmetric* structures described by a *pair* of general, nondecreasing and possibly distinct cost functions: an *earliness* and a *tardiness* cost function which applies to the *early* and *tardy* jobs, respectively. This heuristic is developed in §4 and covers the general class of cost structures treated in Kahlbacher (1992) and Federgruen and Mosheiov (1993) for *single* machine problems. See the latter for a discussion of many important, often asymmetric, earliness and tardiness costs. No efficiently computable and tight lower bounds or practical exact solution methods are available for this general class of cost structures. In the single machine case, we are however able to show that the heuristic performs excellently by comparing it against the cost of optimal schedules, computed with the exact method in Federgruen and Mosheiov.

## 1. The Basic Multimachine Model: An $O(N \log N)$ Lower Bound and Heuristic

Consider a scheduling problem with $N$ jobs and $m$ identical machines. All jobs share a common nonrestrictive due date $d$. The cost incurred for job $j$ depends on the absolute value of the deviation $\Delta_j$ of its completion time $C_j$ from the due date $d$, according to a general, nondecreasing, and convex function $F(\cdot)$, i.e., the cost for job $j$ is given by $F(\Delta_j)$, $j = 1, \ldots, N$. The cost structure in the basic model is thus symmetric and convex. For $i = 1, \ldots, N$, let $P_i$ denote the integer-valued processing time for job $i$, and $s_j = \sum_{i=1}^{j} P_i$ the cumulative processing time for the $j$ smallest jobs. The jobs are numbered such that $P_1 \leq \cdots \leq P_N$. The objective is to minimize $\sum_{j=1}^{N} F(\Delta_j)$, the *sum* of the costs incurred for all jobs.

We start with the derivation of a simple but accurate lower bound for the minimum cost value $z^*$. To derive the latter, we first need the following lemma: It is possible, without loss of optimality, to restrict oneself to schedules without idle times between jobs. (This property is easily verified and well known for $m = 1$.) Let $\Delta_{(i)}$ = the $i$th smallest completion time deviation (among all $N$ jobs on all $m$ machines), $i = 1, \ldots, N$.

LEMMA 1. *For any schedule without idle times,*

$$\Delta_{(j+1)} + \Delta_{(j+2)} + \cdots + \Delta_{(j+2m)} \geq s_{j+m},$$
$$j = 0, 1, \ldots, N - 2m.$$

PROOF. Let $\Theta_{(i)}$ denote the set of completion times with the $i$ smallest deviations from the due date (on any machine). Let $\Delta^E_{(i,k)}(\Delta^T_{(i,k)})$ denote the deviation from the due date of the earliest (latest) completion time in $\Theta_{(i)}$ on machine $k$. Observe that

$$\sum_{k=1}^{m} \Delta^E_{(i,k)} + \Delta^T_{(i,k)} \geq s_{i-m}, \quad \text{for } i > m, \tag{1}$$

since the left-hand side represents the sum of the processing times of all jobs in $\Theta_{(i)}$ except for the $m$ jobs that initiate the schedules on the $m$ machines, i.e., the sum of $(i - m)$ processing times. Thus, by (1),

$$\Delta_{(j+1)} + \Delta_{(j+2)} + \cdots + \Delta_{(j+2m)}$$
$$\geq \sum_{k=1}^{m} [\Delta^E_{(j+2m,k)} + \Delta^T_{(j+2m,k)}] \geq s_{j+m}.$$

(The left-hand side of the first inequality represents the sum of the $2m$ *largest* $\Delta$-values in $\Theta_{(j+2m)}$, while its right-hand side represents the sum of $2m$ $\Delta$-values in $\Theta_{(j+2m)}$ as well.) □

We now derive a lower bound for $z^*$ by relaxing the scheduling problem to a mathematical program of simple structure which allows for a closed form expression:

$$\text{(LB)} \quad \underline{z} = \min \sum_{j=1}^{N} F(\Delta_{(j)}), \tag{2}$$

$$\text{s.t.} \quad \sum_{k=1}^{2m} \Delta_{(N-2lm+k)} \geq s_{N-(2l-1)m};$$

$$l = 1, 2, \ldots, \lfloor N/2m \rfloor, \tag{3}$$

$$\Delta_{(i+1)} \geq \Delta_{(i)}; \qquad i = 1, 2, \ldots, N-1; \quad \Delta_{(1)} \geq 0. \tag{4}$$

THEOREM 2 (*Lower Bound*).

$$\underline{z} = \sum_{l=1}^{\lfloor N/2m \rfloor} 2mF(s_{N-(2l-1)m}/2m) \leq z^*. \tag{5}$$

PROOF. The $\Delta$-values satisfy (3) in view of Lemma 1 with $j = N - 2lm$. Constraints (4) follow from the definition of the $\Delta$-values. The mathematical program (LB) is thus a relaxation of the scheduling problem; hence $\underline{z} \leq z^*$. The closed form expression for **z** is verified as follows: note that the vector $\Delta^*$ with

$$\Delta^*_{(1)} = \cdots = \Delta^*_{(N-2m\lfloor N/2m \rfloor)} = 0,$$
$$\Delta^*_{(N-2lm+1)} = \Delta^*_{(N-2lm+2)} = \cdots$$
$$= \Delta^*_{(N-2lm+2m)} = s_{N-(2l-1)m}/2m$$
$$\text{for all } l = 1, \ldots, \lfloor N/2m \rfloor, \tag{6}$$

minimizes (2) subject to (3). This holds since in the latter optimization problem $\Delta_{(1)}, \ldots, \Delta_{(N-2m\lfloor N/2m \rfloor)}$ are unconstrained and since $F(\cdot)$ is nondecreasing while the remaining problem decomposes into $\lfloor N/2m \rfloor$ separate *single* constraint problems, the $l$th of which consists of minimizing $\sum_{k=1}^{2m} F(\Delta_{(N-2lm+k)})$ subject to the $l$th constraint in (3). Since $F$ is convex, the optimum solution for this single constraint problem is given by the values in (6).

The $\Delta^*$-vector thus minimizes (2) subject to (3). Since it satisfies (4) as well, it is optimal for the complete mathematical program (LB). Substitution of the $\Delta^*$-values in (2) results in the expression for $\underline{z}$. □

The lower bound expression in (5) can be simplified in the following special cases:

(a) $N = 2rm$ for some $r \geq 1$: $\underline{z} = 2m \sum_{l=1}^{r} F(s_{(2l-1)m}/2m)$.

(b) $N = (2r+1)m$ for some $r \geq 1$: $\underline{z} = 2m \sum_{l=1}^{r} F(s_{2lm}/2m)$.

(c) $m = 1$ and $N$ even: $\underline{z} = 2 \sum_{l=1}^{N/2} F(s_{2l-1}/2)$.

(d) $m = 1$ and $N$ odd: $\underline{z} = 2 \sum_{l=1}^{\lfloor N/2 \rfloor} F(s_{2l}/2)$.

The bound in Theorem 2 is *tight*, i.e., $\underline{z} = z^*$, for the following "perfectly symmetric" problem instances.

DEFINITION 1. If $N = 2rm$ for some $r \geq 1$, a problem instance is *perfectly symmetric* if $P_1 = P_2 = \cdots = P_m$; $P_{m+1} = P_{m+2} = \cdots = P_{3m}$; $P_{3m+1} = P_{3m+2} = \cdots = P_{5m}$; $\ldots$; $P_{(2r-3)m+1} = P_{(2r-3)m+2} = \cdots = P_{(2r-1)m}$.

DEFINITION 2. If $N = (2r+1)m$ for some $r \geq 1$, a problem instance is *perfectly symmetric* if $P_1 = P_2 = \cdots = P_{2m}$; $P_{2m+1} = P_{2m+2} = \cdots = P_{4m}$; $P_{4m+1} = P_{4m+2} = \cdots = P_{6m}$; $\ldots$; $P_{2(r-1)m+1} = P_{2(r-1)m+2} = \cdots = P_{2rm}$.

Tightness of the lower bound for perfectly symmetric instances is verified by considering the following simple *Alternating Schedule*. (This is a generalization of a schedule introduced by Eilon and Chowdhury 1977 for minimizing the variance of waiting times on a single machine. In specifying this Alternating Schedule, we assume, without loss of generality, that $N$ is a multiple of $m$; if $N$ fails to be a multiple of $m$, increase the number of jobs to $\lceil N/m \rceil m$, by adding $(m\lceil N/m \rceil - N)$ dummy jobs of zero processing time.)

DEFINITION 3. The Alternating Schedule schedules the jobs in consecutive batches of size $m$, i.e., first jobs $\{1, \ldots, m\}$ are scheduled, then jobs $\{m + 1, \ldots, 2m\}$, etc. The $i$th job in each batch is scheduled on machine $i$ $(i = 1, \ldots, m)$. If $N = (2r + 1)m$ for some $r \geq 0$, the heuristic schedules the first batch to be *completed* at the due date. The remaining batches of jobs are scheduled alternatingly at the tail and head of the partial schedules constructed thus far, starting with the second batch, which is scheduled right *after* the first one. If $N = 2rm$ for some $r \geq 1$, the Alternating Schedule schedules the first batch so that exactly *half* of each job in the batch is completed at the due date. The remaining batches of jobs are scheduled alternatingly at the head and tail of the partial schedules constructed thus far, starting with the second batch which is scheduled *in front of* the first one.

Let $z^{AS}$ denote the cost of the schedules generated by the Alternating Schedule.

PROPOSITION 3. $z^{AS} = \underline{z} = z^*$ *for any perfectly symmetric instance. In other words, the Alternating Schedule is optimal and the lower bound* $\underline{z}$ *tight for any perfectly symmetric instance.*

PROOF. Consider a perfectly symmetric instance with $N = (2r + 1)m$ for some $r \geq 0$. (The proof for the case $N = 2rm$ is similar; see Federgruen and Mosheiov 1994). Observe that all $m$ machines have identical schedules. Since the instance is perfectly symmetric, the cost of the schedule for machine $i$ is given by:

$$\begin{aligned}&\{F(P_i) + F(P_{m+i})\} + \{F(P_i + P_{2m+i}) + F(P_{m+i} + P_{3m+i})\}\\ &\quad + \cdots + \{F(P_i + P_{2m+i} + \cdots + P_{2(r-1)m+i})\\ &\quad + F(P_{m+i} + P_{3m+i} + \cdots + P_{2(r-1)m+m+i})\}\\ &= 2F(s_{2m}/2m) + 2F(s_{4m}/2m) + \cdots + 2F(s_{2rm}/2m).\end{aligned}$$

The cost of all schedules is thus given by

$$z^{AS} = 2m \sum_{l=1}^{r} F(s_{2lm}/2m) = \underline{z}. \quad \square$$

Both the lower bound and the Alternating Schedule heuristic require no more than $O(N \log N)$ elementary operations. Moreover, the Alternating Schedule is independent of the specific cost function $F$ and is therefore insensitive to changes in the cost parameters (as long as a symmetric cost structure is maintained). Also, the job sequences on the $m$ machines merely depend on the *ordinal* ranking of the processing times rather than their *cardinal* values.

Most problem instances are, of course, not perfectly symmetric, but the tightness (optimality) of the lower bound (heuristic) for perfectly symmetric instances bodes well for the bound (heuristic)'s general performance. Assume, for example, that $N = (2r + 1)m$. Any instance which fails to be perfectly symmetric may be approximated by a perfectly symmetric instance by replacing the processing times in each consecutive block of $2m$ jobs by the average processing time in that block, i.e.,

$$P_{2lm+j} := \frac{1}{2m} \sum_{i=1}^{2m} P_{2lm+i} \quad \text{for all}$$

$$j = 1, \ldots, 2m \ (1 = 0, \ldots, r).$$

(In particular, when $N$ is large and all processing times belong to a bounded interval, these adjustments are relatively minor.) Note, by Proposition 3, that

$$\underline{z} = 2m \sum_{l=1}^{r} F(s_{2lm}/2m),$$

the lower bound for the original instance, equals both the *minimum* cost value and the cost of the Alternating Schedule in the approximating (perfectly symmetric) instance; moreover, the job sequences generated by the Alternating Schedule for the original instance are identical to those generated for the approximating instance, since the *relative* ranking of the processing times is unaffected by the transformation. The next two sections provide a rigorous foundation for the above intuition regarding the performance of the lower bound and Alternating Schedule.

## 2. Worst Case Performance Analysis of Lower Bound and Alternating Schedule

In this section we show that the gap between the cost of the Alternating Schedule and the lower bound is bounded by an expression which rapidly decreases to zero as $N$ increases to infinity, if the processing times are uniformly bounded or if they are i.i.d. with finite mean and variance. In other words, the Alternating Schedule is asymptotically optimal, and the lower bound is asymptotically accurate. The optimality and accuracy gaps depend on $\sigma$, a measure of asymmetry in the processing times defined as follows:

$$\sigma = \sum_{i=1}^{r} \sigma_i \quad \text{where}$$

if $N = (2r + 1)m$:

$$\sigma_i = \max\left\{ P_{2im} - \frac{1}{2m}(s_{2im} - s_{2(i-1)m}), \frac{1}{2m}(s_{2im} - s_{2(i-1)m}) - P_{2(i-1)m+1} \right\}; \quad i = 1, \ldots, r,$$

if $N = 2rm$: $\sigma_1 = \max\{P_m - s_m/m; s_m/m - P_1\}$;

$$\sigma_i = \max\left\{ P_{(2i-1)m} - \frac{1}{2m}(s_{(2i-1)m} - s_{(2(i-1)-1)m}); \frac{1}{2m}(s_{(2i-1)m} - s_{(2(i-1)-1)m}) - P_{(2i-3)m+1} \right\}; \quad \text{for } i = 2, \ldots, r.$$

Note that $\sigma \geq 0$ and $\sigma = 0$ for perfectly symmetric instances. We also obtain a worst case bound for the Alternating Schedule's optimality gap, and the lower bound's accuracy gap, which applies to *all* values of $N$, is independent of $N$ and increases slowly with $\sigma$. If the function $F(X) \sim X^q$ as $X \to \infty$ for some $q > 1$ (see condition $C_2$ below), and a minor additional technical condition holds, this worst case gap is in fact uniformly bounded in all problem parameters (including $\sigma$). We give separate treatments to the cases where $N$ is an odd or even multiple of $m$. Let $X_i = \sum_{l=1}^{i} \sigma_l$. Clearly $0 \leq X_1 \leq X_2 \leq \cdots \leq X_r = \sigma$.

LEMMA 4. *(a) Let $N = (2r + 1)m$ for some $r \geq 0$ and fix $\sigma$ and the values $s_{2m}, s_{4m}, \ldots, s_{2rm}$.*

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \leq \frac{\sum_{i=1}^{r} \{F(s_{2im}/2m + \sigma) + F(s_{2im}/2m - \sigma) - 2F(s_{2im}/2m)\}}{2 \sum_{i=1}^{r} F(s_{2im}/2m)}.$$

*(b) Let $N = 2rm$ for some $r \geq 1$, and fix $\sigma$ and the values $s_m, s_{3m}, s_{5m}, \ldots, s_{(2r-1)m}$*

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \leq \frac{\sum_{i=0}^{r-1} \{F(s_{(2i+1)m}/2m + \sigma) + F(s_{(2i+1)m}/2m - \sigma) - 2F(s_{(2i+1)m}/2m)\}}{2 \sum_{i=0}^{r-1} F(s_{(2i+1)m}/2m)}.$$

PROOF. (a) Recall that the Alternating Schedule schedules jobs in batches of size $m$. The $m$ smallest jobs are all completed at the due date and incur zero cost. For $i = 1, \ldots, r$ consider now the completion time deviations for the $2i$th and $(2i + 1)$st batch of jobs. Each of these deviations equals the sum of $i$ processing times, one from the range $[P_1, \ldots, P_{2m}]$, one from the range $[P_{2m+1}, \ldots, P_{4m}]$, etc, with the $i$th processing time taken from the range $[P_{2(i-1)m+1}, \ldots, P_{2im}]$. The cost incurred under the Alternating Schedule, for the $2m$ jobs in the $2i$th and $(2i + 1)$st batch can thus be represented as

$$\sum_{l=1}^{2m} F(s_{2im}/2m + y_l), \quad \text{where} \tag{7}$$

$$\sum_{l=1}^{2m} y_l = 0 \quad \text{and} \quad |y_l| \leq X_i \text{ for all } l = 1, \ldots, 2m. \tag{8}$$

To obtain an upper bound for the expression in (7) we maximize this expression over the simplex defined by (8). Since $F$ is convex, this maximum is achieved in one of the extreme points of the simplex. Also, since (9) is symmetric in $(y_1, \ldots, y_{2m})$, the maximum is achieved by setting $m$ of the $y$-variables equal to $+X_i$, and $m$ of them equal to $-X_i$. The cost incurred under the Alternating Schedule for the jobs in the $2i$th and $(2i + 1)$st batch is thus bounded by

$$m[F(s_{2im}/2m + X_i) + F(s_{2im}/2m - X_i)]$$

and the total cost over all jobs by

$$\sum_{i=1}^{r} m[F(s_{2im}/2m + X_i) + F(s_{2im}/2m - X_i)]$$

$$\leq m \sum_{i=1}^{r} [F(s_{2im}/2m + \sigma) + F(s_{2im}/2m - \sigma)],$$

where the inequality follows from the left hand side being nondecreasing in $X$ in view of the convexity of $F(\cdot)$.

(b) The proof of part (b) is analogous to that of part (a). □

Lemma 4 allows us to prove that the Alternating Schedule has a worst case optimality gap which is bounded by a simple function of the asymmetry value $\sigma$, provided the function $F(\cdot)$ is log-concave, and continuously differentiable, i.e.,

*Condition* $(C_1)$. $\ln F(\cdot)$ is concave and $F(\cdot)$ is continuously differentiable with derivative $F'(\cdot)$.

THEOREM 5. *Assume $F(\cdot)$ is convex and satisfies condition $(C_1)$.*

*(a) Let $N = (2r+1)m$ for some $r \geq 0$, and fix $\sigma$ and the values $s_{2m}, s_{4m}, \ldots, s_{2rm}$. Then,*

$$\frac{z^{AS} - z^*}{z^*} \leq \frac{\sigma}{2} \frac{F'(s_{2m}/2m + \sigma)}{F(s_{2m}/2m)}. \tag{9}$$

*(b) Let $N = 2rm$ for some $r \geq 1$, and fix $\sigma$ and the values $s_m, s_{3m}, s_{5m}, \ldots, s_{(2r-1)m}$. Then,*

$$\frac{z^{AS} - z^*}{z^*} \leq \frac{\sigma}{2} \frac{F'(s_m/2m + \sigma)}{F(s_m/2m)}. \tag{10}$$

PROOF. (a) It follows from Lemma 4 that

$$\begin{aligned} z^{AS} - \underline{z} &\leq \sum_{i=1}^{r} m\{F(s_{2im}/2m + \sigma) \\ &\qquad + F(s_{2im}/2m - \sigma) - 2F(s_{2im}/2m)\} \\ &= \sum_{i=1}^{r} m[F(s_{2im}/2m + \sigma) - F(s_{2im}/2m)] \\ &\quad + \sum_{i=1}^{r} m[F(s_{2im}/2m - \sigma) - F(s_{2im}/2m)] \\ &\leq \sum_{i=1}^{r} m[F(s_{2im}/2m + \sigma) - F(s_{2im}/2m)] \\ &\leq \sigma m \sum_{i=1}^{r} F'(s_{2im}/2m + \sigma), \end{aligned} \tag{11}$$

where the first inequality follows from $F(\cdot)$ being nondecreasing and the second inequality follows from $F(\cdot)$ being convex. Observe that the function

$$h(x) \stackrel{\text{def}}{=} F'(x+\sigma)/F(x)$$

is nonincreasing in $x$. This may be verified by representing $h(x) = h_1(x)h_2(x)$ where

$$h_1(x) \stackrel{\text{def}}{=} \frac{d \ln F(x+\sigma)}{dx} = \frac{F'(x+\sigma)}{F(x+\sigma)} \geq 0$$

is nonincreasing by the log-concavity of $F$, and

$$h_2(x) \stackrel{\text{def}}{=} F(x+\sigma)/F(x)$$

is nonincreasing since

$$\ln[F(x+\sigma)/F(x)] = \ln F(x+\sigma) - \ln F(x)$$

is nonincreasing, again by the log-concavity of $F$. We thus obtain:

$$\begin{aligned} \frac{z^{AS} - \underline{z}}{\underline{z}} &\leq \frac{\sigma}{2} \sum_{i=1}^{r} \left\{ \frac{F(s_{2im}/2m)}{\sum_{i=1}^{r} F(s_{2im}/2m)} \right\} h(s_{2im}/2m) \\ &\leq \frac{\sigma}{2} \max_{i=1,\ldots,r} \{h(s_{2im}/2m)\} \leq \frac{\sigma}{2} h(s_{2m}/2m) \\ &= \frac{\sigma}{2} F'(s_{2m}/2m + \sigma)/F(s_{2m}/2m). \end{aligned}$$

(b) The proof of part (b) is analogous to that of part (a). □

Observe that the worst case bounds in (9)–(10) decrease to zero as $\sigma \downarrow 0$. A constant worst case optimality gap (independent of $\sigma$) and asymptotic optimality are obtained under the following condition:

*Condition* $(C_2)$. $F(\cdot)$ is continuously differentiable, log-concave and $F(x) \sim x^q$, as $x \to \infty$ for some $q > 1$, i.e., there exists a constant $c > 0$ such that

$$\lim_{x \to \infty} \frac{F(x)}{x^q} = c \quad \text{for some } q > 1.$$

Under condition $(C_2)$, let

$$\phi(x) = \int_0^x F(y)dy \sim x^{q+1} \quad \text{as } x \to \infty.$$

(Since $F$ is differentiable, it is clearly integrable.) In the remainder of this section we assume $N = (2r+1)m$ for some $r \geq 1$; similar proofs apply for general values of $N$.

THEOREM 6. *Assume $F$ satisfies condition $(C_2)$.*

*(a) The Alternating Schedule has a constant worst case optimality gap.*

(*b*) *Further assume* $\sigma(\leq P_N)$ *is uniformly bounded in N. Then,* $\lim_{N\to\infty}(z^{AS} - \underline{z})/\underline{z} = 0$, *and the convergence rate is* $O(N^{-1})$.

PROOF. Note, from the monotonicity of the *P*-values, that $s_{2im}/2m \geq is_{2m}/2m$ for all $i = 1, \ldots, r$. Thus, for $\alpha = s_{2m}/2m$,

$$\underline{z} = 2mF(s_{2im}/2m) \geq 2m\left[F(s_{2rm}/2m) + \sum_{i=1}^{r-1} F(i\alpha)\right].$$

Also, $h(i) \stackrel{\text{def}}{=} F(i\alpha)$ is a nondecreasing function of $i$. Thus, $F(i\alpha) \geq \int_{i-1}^{i} F(\alpha x)dx$ and

$$\begin{aligned}\underline{z}/2m &\geq F(s_{2rm}/2m) + \sum_{i=1}^{r=1} F(i\alpha)\\ &\geq F(s_{2rm}/2m) + \int_0^{r-1} F(\alpha x)dx\\ &= F(s_{2rm}/2m) + \alpha^{-1}\phi((r-1)\alpha). \qquad (12)\end{aligned}$$

(a) Let $\beta = s_{2rm}/2m$. It follows from (11) that:

$$\begin{aligned}\frac{z^{AS} - \underline{z}}{\underline{z}} &\leq \sigma \sum_{i=1}^{r} F'(s_{2im}/2m + \sigma)/2 \sum_{i=1}^{r} F(s_{2im}/2m)\\ &\leq \frac{\sigma}{2} rF'(s_{2rm}/2m + \sigma)\\ &\quad /[F(s_{2rm}/2m) + \alpha^{-1}\phi((r-1)\alpha)]\\ &\leq \tfrac{1}{2}\beta NF'(\beta(2m+1))/[F(\beta) + \alpha^{-1}\phi((r-1)\alpha)],\end{aligned}$$

by (12), since $2m\beta = s_{2rm} \geq P_{2rm} \geq \sigma$ and since $F'(\cdot)$ is nondecreasing. (Note that (11) applies whenever $F$ is convex and differentiable.) By condition $(C_2)$ the upper bound is continuous in $\beta$ and $N$, converges to a constant as $\beta \to \infty$ and to 0 as $N \to \infty$, and is therefore uniformly bounded in $\beta$ and $N$.

(b) It follows from the proof of Lemma 4 that:

$$\begin{aligned}z^{AS} - \underline{z} &\leq \sum_{i=1}^{r} m[F(s_{2im}/2m + X_i) + F(s_{2im}/2m - X_i)]\\ &\quad - 2m\sum_{i=1}^{r} F(s_{2im}/2m)\\ &\leq m\sum_{i=1}^{r}[F(s_{2im}/2m + X_i) - F(s_{2im}/2m)]. \qquad (13)\end{aligned}$$

Recall that $0 \leq X_1 \leq X_2 \leq \cdots \leq X_r \leq \sigma$. For a fixed choice of $\sigma$, and given these constraints with respect to the $X$-variables, the right-hand side of (13) is maximized for $\sigma_1 = X_1 = X_2 = \cdots = X_r = \sigma$. But in that case,

$$X_i = \sigma_1 \leq P_{2m} - P_1 \leq P_{2m} \leq (s_{2(i+1)m} - s_{2im})/2m$$

for all $i = 1, 2, \ldots, r$ where the last inequality follows from the jobs being numbered in nondecreasing order of their processing times. Since $F(\cdot)$ is nondecreasing, the upper bounds for $X_i$ may be used in conjunction with (13) to conclude:

$$\begin{aligned}&z^{AS} - \underline{z}\\ &\leq m\left[\sum_{i=1}^{r} F((s_{2im} + s_{2(i+1)m} - s_{2im})/2m) - F(s_{2im}/2m)\right]\\ &= m[F(s_{2(r+1)m}/2m) - F(s_{2m}/2m)] \leq mF(s_{2rm}/2m).\end{aligned}$$

By (12),

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \leq \frac{\alpha}{2} F\left(\sum_{i=1}^{N} P_i/2m\right)\Big/\phi(rs_{2m}/2m). \qquad (14)$$

This implies that $(z^{AS} - \underline{z})/\underline{z} \to 0$ as $N \to \infty$, with a convergence rate of $O(\frac{1}{N})$. □

Alternatively, we show that the Alternating Schedule is *almost surely* asymptotically optimal when the processing times are independent and identically distributed with a common general distribution, possibly with *unbounded* support.

COROLLARY 7. *Assume F is convex and satisfies condition* $(C_2)$. *Let* $P_1, P_2, \cdots$ *be i.i.d. with a common distribution with finite mean* $\mu$ *and standard deviation* $\sigma$. *Then,*

$$\lim_{N\to\infty} \frac{z^{AS} - z^*}{z^*} = \lim_{N\to\infty} \frac{z^* - \underline{z}}{\underline{z}} = 0$$

*almost surely, where the rate of convergence is* $O(\frac{1}{N})$.

PROOF. Fix $\epsilon > 0$, arbitrarily small. It follows from the central limit theorem that almost surely $\Sigma_{i=1}^{N} P_i \leq N(\mu + \epsilon)$ for all $N$ sufficiently large. It thus follows from (14) that *almost surely*

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \leq \frac{\alpha}{2} F(N(\mu + \epsilon)/2m)/\phi(\lfloor N/2m \rfloor s_{2m}/2m).$$

The remainder of the proof is identical to that of Theorem 6. □

REMARK 1. It follows from the proofs of Lemma 4 and Theorem 2 that the Alternating Schedule continues to be asymptotically optimal and the lower bound continues to be asymptotically accurate when the processing times are uniformly bounded from below and when $F(\cdot)$ satisfies condition $(C_2)$ and is *asymptotically* convex only, i.e., there exists a constant $M > 0$ such that $F(x)$ is convex for $x \geq M$. See Federgruen and Mosheiov (1994) for details.

A faster rate of convergence of the optimality accuracy gaps can be shown under a slightly stronger condition than $(C_2)$:

*Condition* $(C_2')$. $F$ is twice differentiable, $F' \geq 0$, $F'' \geq 0$, $F(x) \sim x^q$ for some $q > 1$ and $F''(\cdot)$ is asymptotically monotone.

THEOREM 8. *Assume $F(\cdot)$ satisfies condition $(C_2')$ and $\sigma(N)(\leq P_N)$ is uniformly bounded in $N$. Then,*

$$\frac{z^{AS} - z^*}{z^*} = O\left(\frac{1}{N^2}\right), \qquad N \to \infty.$$

PROOF. see appendix.

REMARK 2. As with Theorem 6, it is again possible to extend the result in Theorem 8 to an *almost sure* $O(\frac{1}{N^2})$ convergence rate in a probabilistic model in which all processing times are i.i.d. with a common distribution with finite mean and variance.

# 3. Minimizing the Completion Time Variance

In this section we discuss the special case where the *variance* of the completion times is to be minimized. This specific objective has received a great deal of attention, all of which has been confined to the *single* machine problem, i.e., the case where $m = 1$.

The problem was first introduced in Merten and Muller (1972) in the context of file organization procedures for which it is important to provide balanced response times. Schrage (1975) derived a number of structural properties and an algorithm for scheduling up to five jobs. Hall and Kubiak (1991) recently proved a conjecture of Schrage regarding the position of the second and third largest job in an optimal schedule. Eilon and Chowdhury (1977) and Bagchi et al. (1987) developed enumeration schemes which can be comfortably used to find an optimal schedule when the number of jobs is small (say, $N \leq 20$). Heuristics, again restricted to the single machine case and without worst-case or probabilistic characterizations of optimality gaps, are due to Kanet (1981), Vani and Raghavachari (1987), and Gupta et al. (1990). De et al. (1993) developed an exact but pseudo-polynomial algorithm. Kubiak (1993a) and Jozefowska and Kubiak (1993) provide alternative approaches to the single machine variance minimization problem. The latter solve problems with up to 50 jobs; their three heuristics are fairly accurate, compared to earlier approaches, but their complexity is $O(N^4)$ or $O(N^3)$.

Indeed, Kubiak (1993b) proved that the problem is binary NP-hard (even when $m = 1$). Ventura and Weng (1993) obtain an accurate *lower bound* by solving the Lagrangian dual of a new mixed integer programming formulation. The required CPU-time grows quickly with $N$; a FORTRAN code run on a VAX-8550 machine requires, for instances with $N = 500$ jobs, an average in excess of 10,000 seconds.

Ours appear to be the first lower bound and heuristic for the *multimachine* variance minimization problem. We first prove that the problem is equivalent to that of minimizing the sum of squared deviations of the jobs' completion times from a (sufficiently large) due date $d$, i.e., the special case where $F(x) = x^2$. (Bagchi et al. 1987 established this result for $m = 1$.) Let $z_d^*$ ($z_{\text{var}}^*$) denote the minimum value of (2) with $F(x) = x^2$ (the minimum completion time variance).

PROPOSITION 9. *An optimal schedule for problem* (2) *with $F(x) = x^2$ is also optimal when minimizing the variance of the completion times. Moreover, $z_d^* = z_{\text{var}}^*$.*

PROOF. For any given schedule let $z_d(\pi)(z_{\text{var}}(\pi))$ denote the value of (2) with $F(x) = x^2$ (the completion time variance) under this schedule. For both objectives, it is optimal to process the jobs that are assigned to a given machine, without intermittent idle times. Thus, for any given variance minimizing schedule $\pi^*$, we obtain a *different* variance minimizing schedule by postponing the starting time of each of the machines' sequences by a constant $c$. (The mean completion time of the translated schedule is then augmented by $c$ as well.) In other words, there exists a variance minimizing schedule $\pi_{\text{var}}^*$ for *any* prespecified mean completion time $\bar{C}$ sufficiently large, e.g., $\bar{C} \geq s_N$. Thus, choose $\bar{C} = d^0 \geq s_N$, and let $\pi^*_{d^0}$ be an optimal schedule for (2) with $d = d^0$.

Since for a given set of completion times $\{C_j: j = 1, \ldots, N\}$ $\sum_{j=1}^N (C_j - \bar{C})^2 \leq \sum_{j=1}^N (C_j - d)^2$ for *all* $d$, we obtain

$$z^*_{\text{var}} = z_{\text{var}}(\pi^*_{\text{var}}) \leq z_{\text{var}}(\pi^*_{d^0}) \leq z_{d^0}(\pi^*_{d^0})$$

$$= z^*_{d^0} \leq z_{d^0}(\pi^*_{\text{var}}) = z_{\text{var}}(\pi^*_{\text{var}}) = z^*_{\text{var}}. \quad \square$$

Observe that the function $F$, with $F(x) = x^2$, satisfies each of the conditions $(C_1)$, $(C_2)$, and $(C_2')$. We thus conclude:

COROLLARY 10. *Let $z^{AS}_{\text{var}}$ denote the variance of the completion times in the Alternating Schedule. Let*

$$\underline{z} = (2m)^{-1} \sum_{l=1}^{\lfloor N/2m \rfloor} s^2_{N-(2l-1)m} \quad \textit{and} \quad \alpha = s_{2m}/2m.$$

(*a*)

$$\frac{z^{AS}_{\text{var}} - \underline{z}}{\underline{z}} \leq \frac{\sigma^2 + \alpha\sigma}{\alpha^2}.$$

(*b*) *Fix N. The optimality gap of $z^{AS}$ and the accuracy gap of the lower bound $\underline{z}$ can both be bounded by a constant (which is independent of the processing times).*

(*c*) *When the processing times are uniformly bounded,*

$$\frac{z^{AS}_{\text{var}} - \underline{z}}{\underline{z}} = O\left(\frac{1}{N^2}\right), \qquad N \to \infty.$$

*The same limit results hold almost surely when the processing times are i.i.d. with finite mean and variance.*

PROOF. Let $z^{AS}$ denote the value of $\sum_j F(\Delta_j)$ with $F(x) = x^2$ obtained by the Alternating Schedule. Clearly $z^{AS}_{\text{var}} \leq z^{AS}$. Also, $z^*_{\text{var}} = z^*_d$ by Proposition 9. Thus,

$$(z^{AS}_{\text{var}} - z^*_{\text{var}})/z^*_{\text{var}} \leq (z^{AS} - z^*_d)/z^*_d \quad \text{and}$$

$$(z^*_{\text{var}} - \underline{z})/\underline{z} = (z^*_d - \underline{z})/\underline{z}.$$

Parts (a)–(c) thus follow from Theorem 5, Theorem 6, and Theorem 8, respectively. $\square$

We have gauged the *empirical* performance of the Alternating Schedule and lower bound by conducting a numerical study with 300 problem instances, partitioned into 12 sets of 25 instances each. All instances have $m = 3$ machines; see Mosheiov (1991) for instances with a single machine. $N$ varies from 30 in set 1 to 3,000 in set 12. The processing times are generated randomly from the uniform distribution on the integers in the interval [1, 100].

Table 1 exhibits for each set the *average* value of the ratios $(z^{AS} - \underline{z})/\underline{z}$, an upper bound for the optimality, and accuracy gap. Our main observation is that these ratios are extremely close to one, even for relatively small size problems. Note that even for problems with $N = 30$ jobs (i.e., an average of 10 jobs per machine) the gap between the cost of the Alternating Schedule and the lower bound is no more than 0.75% (on average); for $N = 300$, the average gap is down to 0.008%; and for $N = 3{,}000$ it is down to 0.0001%. These results confirm those of Corollary 10, in particular asymptotic optimality with an $O(\frac{1}{N^2})$ convergence rate: the average gap (GAP) is reduced by a factor of 100 (approximately), when the number of jobs is increased by a factor of 10. Indeed, regressing GAP against the variable $N^{-2}$ we obtain the regression equation GAP = $6.83702N^{-2}$ with an $R^2$-value of 0.999991! (The standard deviation of the estimate of the proportionality constant is $7 \times 10^{-6}$. Addition of a term proportional with $N^{-1}$ does not improve the fit significantly.)

Recall that the *complexity* of the bound and heuristic is $O(N \log N)$ only. (After the initial sorting of the processing times, all remaining work is in fact *linear* in $N$.) This complexity measure is confirmed by the CPU times observed with a FORTRAN code run on an IBM 4381 (VM/CS): an average CPU time of 0.04 and 0.19 seconds is required for instances with $N = 300$ and $N = 3{,}000$, respectively. We conclude that for all practical purposes, the Alternating Schedule and lower bound can be used to very quickly generate a close-to-optimal solution and (ex post) bound for the optimality gap. This applies even when the number of jobs is small, or in the single machine case where reasonably efficient

**Table 1** **The Completion Time Variance Problem**

| $N$ | $(z^{AS} - \underline{z})/\underline{z}$ | $N$ | $(z^{AS} - \underline{z})/\underline{z}$ |
|---|---|---|---|
| 30 | 0.0075981 | 1500 | 0.00000335 |
| 90 | 0.0008232 | 1800 | 0.00000227 |
| 300 | 0.0000820 | 2100 | 0.00000190 |
| 600 | 0.0000209 | 2400 | 0.00000158 |
| 900 | 0.00000915 | 2700 | 0.00000124 |
| 1200 | 0.00000483 | 3000 | 0.00000104 |

solution methods can be invoked, e.g., those of Kahlbacher (1992) and Federgruen and Mosheiov (1993).

## 4. Asymmetric General Cost Structures

The basic model treated in the previous sections assumes a symmetric and convex cost structure. In many settings one finds, however, that the cost structure is *asymmetric*, i.e., it is specified by *two* distinct, nonnegative and nondecreasing functions $F(\cdot)$ and $G(\cdot)$, where $F(E)$ $[G(T)]$ denotes the cost of a job being early (tardy) by $E(T)$ time units ($F(0) = G(0) = 0$.) Moreover, $F$ or $G$ may fail to be convex. Asymmetric cost structures arise, e.g., when the due date is restrictive, i.e., $F(x) = G(x)$, if $x \leq d$, and $F(x) = \infty$, otherwise. More generally, one finds that earliness costs arise from inventory carrying and maintenance costs, while tardiness costs relate to explicit or contractually agreed upon lateness penalties or estimates of (implicit) goodwill losses. The *shapes* of the earliness and tardiness cost functions $F$ are therefore often quite different. See Federgruen and Mosheiov (1993) for a more detailed discussion and specific examples.

The strong performance of the Alternating Schedule for symmetric cost structures, suggests the following Modified Alternating Scheduling Heuristic (MASH) for general *asymmetric* structures. As before, we assume that $N$ is a multiple of $m$.

***(MASH)***

*Case* 1. $N = (2r + 1)m$ for some $r \geq 0$.

*Step* 1. For $j = 1, \ldots, m$, schedule job $j$ on machine $j$ to be completed at the due date. Next, proceeding in the sequence in which the jobs are numbered, schedule each of the remaining jobs $m + 1, m + 2, \ldots, N$ either directly preceding or directly following one of the machines' partially constructed schedules, wherever (among the $2m$ possible choices) the resulting cost for this job is minimal.

*Step* 2. Shift each of the machines' schedules by varying its starting time to its optimal value.

*Case* 2. $N = 2rm$ for some $r \geq 1$.

*Step* 1. For $j = i, \ldots, m$ schedule job $j$ on machine $j$, to be completed at time $C_j^*$ with $d \leq C_j^* \leq d + P_j$ such that

$$G(C_j^* - d) + F(d - C_j^* + P_j) = \min_x \{G(x - d) + F(d - x + P_j)\}.$$

Schedule the remaining jobs as in Case 1.

*Step* 2. As in Case 1.

Thus, in Step 1 jobs are assigned in a greedy manner.

Observe that (MASH) generates a schedule for each machine which is:

(1) *of V-shape*, i.e., a (possibly empty) sequence of jobs of decreasing length is followed by a (possibly empty) sequence of jobs of increasing length;

(2) *(LPTB/SPTA)*, i.e., the sequence of jobs completed *before* (*after*) the due date has decreasing (increasing) processing times;

(3) without intermittent idle times.

All these properties are known to hold for an optimal schedule for each of the $m$ machines, see, e.g., Federgruen and Mosheiov.

Observe that (MASH) reduces to the basic Alternating Schedule when $F = G$, i.e., under symmetric cost structures. Moreover, the complexity of the (MASH) heuristic continues to amount to $O(N \log N)$ elementary operations and $O(mN)$ evaluations of the $F$- and $G$-functions in Step 1, and $m$ minimizations of nonlinear functions of a single variable in Step 2.

For multiple machines ($m \geq 2$), no efficiently computable and tight lower bounds have been identified for general, asymmetric cost structures. (See, however, Mosheiov 1991, subsection 2.3.2 for lower bounds in case $m = 1$ and the cost structure is symmetric except for a restrictive due date. These bounds are based on the solution of a mathematical program similar to (LB), but fail to be tight.) We are thus unable to gauge (MASH')s performance except in the single machine case, where the dynamic programming methods in Kahlbacher (1992) or Federgruen and Mosheiov (1993) can be used.

For the latter, we have conducted a numerical study with 500 (single machine) instances, partitioned into 20 sets of 25 instances each; see Table 2. In the first ten sets, $F$ and $G$ are both convex; in the first (second) quintuple of sets the tardiness cost function is chosen to increase more (less) rapidly than the (quadratic) earliness cost function. Sets 10–15 have concave cost functions, and the last five sets represent the case of a symmetric (in

**Table 2 Performance of MASH**

| $N$ | $F$ | $G$ | $d$ | OPT | MASH | RATIO |
|---|---|---|---|---|---|---|
| 25 | $x^2$ | $x$ | n.r.[1] | 10,014.8 | 10,434.0 | 1.0419 |
| 50 | $x^2$ | $x$ | n.r. | 40,739.7 | 41,590.1 | 1.0209 |
| 75 | $x^2$ | $x$ | n.r. | 90,731.4 | 91,705.5 | 1.0107 |
| 100 | $x^2$ | $x$ | n.r. | 162,177.8 | 163,256.0 | 1.0067 |
| 200 | $x^2$ | $x$ | n.r. | 649,964.6 | 651,647.7 | 1.0026 |
| 25 | $x^2$ | $x^3$ | n.r. | 5,099,580 | 5,262,392 | 1.0319 |
| 50 | $x^2$ | $x^3$ | n.r. | 45,064,448 | 45,309,488 | 1.0054 |
| 75 | $x^2$ | $x^3$ | n.r. | 154,127,104 | 154,641,760 | 1.0033 |
| 100 | $x^2$ | $x^3$ | n.r. | 377,999,104 | 378,529,024 | 1.0014 |
| 200 | $x^2$ | $x^3$ | n.r. | 1,073,739,260 | 1,102,062,590 | 1.0009 |
| 25 | $10\sqrt{x}$ | $x$ | n.r. | 3,456.9 | 3,482.3 | 1.0075 |
| 50 | $10\sqrt{x}$ | $x$ | n.r. | 10,544.5 | 10,570.5 | 1.0025 |
| 75 | $10\sqrt{x}$ | $x$ | n.r. | 19,867.0 | 19,889.5 | 1.0011 |
| 100 | $10\sqrt{x}$ | $x$ | n.r. | 31,144.1 | 31.166.7 | 1.00072 |
| 200 | $10\sqrt{x}$ | $x$ | n.r. | 90,625.4 | 90,647.4 | 1.0024 |
| 25 | $x^2$ | $x^2$ | 250[2] | 4,800,744 | 4,835,104 | 1.00716 |
| 50 | $x^2$ | $x^2$ | 500[2] | 34,593,760 | 34,705,680 | 1.00324 |
| 75 | $x^2$ | $x^2$ | 750[2] | 109,513,904 | 109,944,832 | 1.00393 |
| 100 | $x^2$ | $x^2$ | 1000[2] | 256,010,944 | 256,427,568 | 1.00163 |
| 200 | $x^2$ | $x^2$ | 2000[2] | 1,916,707,099 | 1,918,662,140 | 1.00102 |

[1] n.r. = nonrestrictive.
[2] $d = 0.2$ (Makespan).

fact, quadratic) structure except for a restrictive due date, chosen as $0.2 \times$ makespan ($=0.2s_N$). All instances in a set have the same value of $N = 25, 50, 75, 100$, or 200. The processing times are again generated from the uniform distribution on the integers from 1 to 100. We report for each set, the average value of the minimum cost-value (OPT), that of (MASH)'s cost value and that of the ratios of the latter and the optimal cost value (RATIO).

We observe that (MASH) generates close-to-optimal schedules across different cost structures and even for instances with a relatively small number of jobs $N$. As for the symmetric case, the optimality gap reduces to zero as the number of jobs is increased.

## Appendix

PROOF OF THEOREM 8. It follows from Lemma 4 that

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \le \frac{\sum_{i=1}^{r} [F(s_{2i_m}/2m + \sigma) + F(s_{2i_m}/2m - \sigma) - 2F(s_{2i_m}/2m)]}{2\sum_{i=1}^{r} F(s_{2i_m}/2m)}. \tag{15}$$

Since $F$ is twice differentiable, by Taylor's formula there exist numbers $\zeta_i$ and $\zeta_i'$ with $0 \le \xi_i, \xi_i' \le \sigma$ such that

$$\begin{aligned}
&F\left(\frac{s_{2i_m}}{2m} + \sigma\right) + F\left(\frac{s_{2i_m}}{2m} - \sigma\right) - 2F\left(\frac{s_{2i_m}}{2m}\right)\\
&\quad = \left[F\left(\frac{s_{2i_m}}{2m} + \sigma\right) - F\left(\frac{s_{2i_m}}{2m}\right)\right] + \left[F\left(\frac{s_{2i_m}}{2m} - \sigma\right) - F\left(\frac{s_{2i_m}}{2m}\right)\right]\\
&\quad = \sigma F'\left(\frac{s_{2i_m}}{2m}\right) + \frac{\sigma^2}{2}F''\left(\frac{s_{2i_m}}{2m} + \zeta_i\right)\\
&\qquad - \sigma F'\left(\frac{s_{2i_m}}{2m}\right) + \frac{\sigma^2}{2}F''\left(\frac{s_{2i_m}}{2m} + \zeta_i'\right)\\
&\quad = \frac{\sigma^2}{2}\left[F''\left(\frac{s_{2i_m}}{2m} + \zeta_i\right) + F''\left(\frac{s_{2i_m}}{2m} + \zeta_i'\right)\right].
\end{aligned}$$

By Condition ($C_2'$), there exists a constant $M > 0$ such that $F''(x)$ is monotone for $x \ge M$. Assume without loss of generality that $F''(x)$ is nondecreasing for $x \ge M$. (The case where $F''$ is nonincreasing can be handled similarly.) We thus rewrite (15) as

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \le \left\{K + \frac{\sigma^2}{2}\sum_{i=i_0+1}^{r} [F''(s_{2i_m}/2m + \zeta_i) + F''(s_{2i_m}/2m - \zeta_i')]\right\} \Big/ \sum_{i=1}^{r} 2F(s_{2i_m}/2m),$$

where $i_0 = \lceil 2m(M + \sigma)/s_{2m} \rceil$ and $K$ denotes the sum of the first $i_0$ terms in the summation in the numerator of (15); $K$ is uniformly bounded in $N$ since all processing times are. Observe that for $i \geq (M + \sigma)/(s_{2m}/2m)$,

$$\frac{s_{2im}}{2m} + \sigma \geq \frac{s_{2im}}{2m} + \zeta_i \geq \frac{s_{2im}}{2m} - \zeta_i' \geq \frac{s_{2im}}{2m} - \sigma \geq i s_{2m}/2m - \sigma \geq M.$$

Thus,

$$\frac{z^{AS} - \underline{z}}{\underline{z}} \leq \left\{ \frac{K}{2} + \frac{\sigma^2}{2} \sum_{i=i_0+1}^{r} F''(iP_{\max} + \sigma) \right\} \Big/ \sum_{i=1}^{r} F(iP_1).$$

Since $h(i) \stackrel{\text{def}}{=} F''(iP_{\max} + \sigma)$ and $\bar{h}(i) \stackrel{\text{def}}{=} F(iP_1)$ are nondecreasing functions for $i \geq i_0 + 1$ and $i \geq 1$, respectively, we may replace the summations by integrals:

$$\frac{z^{AS} - z}{z} \leq \frac{K/2 + (\sigma^2/2) \int_{i_0+1}^{r+1} F''(xP_N + \sigma)dx}{\int_0^{r-1} F(xP_1)dx}$$

$$\leq \frac{K/2 + (\sigma^2/2)/P_{\max} F'((r+1)P_{\max} + \sigma)}{P_1^{-1} \phi((r-1)P_1)}$$

$$= O\left(\frac{1}{N^2}\right) \quad \text{as } N, r \to \infty.$$

## References

Bagchi, U., R. Sullivan, and Y. Chang, "Minimizing Mean Squared Deviation of Completion Time about a Common Due Date," *Management Sci.*, 33 (1987), 894–906.

Baker, K. and G. Scudder, "Sequencing with Earliness and Tardiness Penalties: A Review," *Oper. Res.*, 38, 1 (1990), 22–36.

Blackburn, J., "Time-Base Competition: The Next Battleground in American Manufacturing," Irwin, Homewood, IL, 1991.

De, P., J. Ghosh, and C. Wells, "On the Minimization of Completion Time Variance with a Bio Extension," *Oper. Res.*, 40 (1992), 1148–1155.

——, ——, and ——, "A Note on the Minimization of Mean Squared Deviation of Completion Times about a Common Due Date," *Management Sci.*, 35 (1989), 1143–1147.

Eilon, S. and I. Chowdhury, "Minimizing Waiting-time Variance in the Single Machine Problem," *Management Sci.*, 23 (1977), 567–575.

Emmons, H., "Scheduling to a Common Due Date on Parallel Common Processors," *Naval Res. Logistics Quarterly*, 34 (1987), 803–810.

Federgruen, A. and G. Mosheiov, "Simultaneous Optimization of Efficiency and Performance Balance Measures in Single Machine Scheduling Problems," *Naval Res. Logistics Quarterly*, 40 (1993), 951–970.

—— and ——, "Heuristics for Multi-Machine Scheduling Problems with Earliness and Tardiness Costs," Working Paper, Graduate School of Business, Columbia University, New York (unabridged version of this paper), 1994.

Gupta, M., K. Gupta, and C. Bector, "Minimizing the Flow-time Variance in Single-Machine Systems," *J. Opl. Res. Soc.*, 41 (1990), 767–779.

Hall, N., "Single- and Multiple-processor Models for Minimizing Completion-time Variance," *Naval Res. Logistics Quarterly*, 33 (1986), 49–54.

Hall, R., *Zero Inventories*, Dow Jones-Irwin, Homewood, IL, 1983.

—— and W. Kubiak, "Proof of a Conjecture of Schrage about the Completion Time Variance Problem," *Oper. Res. Letters*, 10 (1991), 467–472.

——, ——, and S. Sethi, "Deviation of Completion Times about a Restrictive Common Due Date," *Oper. Res.*, 39 (1991), 847–856.

Jozefowska, J. and W. Kubiak, "Spectral Algorithms for the Completion Time Variance Problems," Working Paper, Faculty of Business Administration, Memorial University of Newfoundland, St. John, Newfoundland, 1993.

Kanet, J., "Minimizing the Average Deviation of Job Completion-times about a Common Due-date," *Naval Res. Logistics Quarterly*, 28 (1981), 643–651.

Kahlbacher, H., "Scheduling with Monotonous Earliness and Tardiness Penalties," *European J. Oper. Res.*, 64 (1992), 258–277.

Kubiak, W., S. Low, and S. Sethi, "Equivalence of Mean Flow Time Problems and Mean Absolute Deviation Problems," *Oper. Res. Letters*, 9 (1990), 371–374.

——, "Completion Time Variance Minimization on a Single Machine is Difficult," *Oper. Res. Letters*, 14 (1993b), 49–59.

——, "New Results on the Completion Time Variance Minimization," to appear in *Discrete Applied Mathematics*.

Merten, A. and M. Muller, "Variance Minimization in Single Machine Sequencing Problems," *Management Sci.*, 18 (1972), 518–528.

Monden, Y., *Toyota Production System*, Institute of Industrial Engineering Press, Norcross, GA, 1983.

Mosheiov, G., "Scheduling Problems with Earliness and Tardiness Costs; A Just-in-Time Approach," Ph.D. Dissertation, Columbia University, New York, 1991.

Schrage, L., "Minimizing the Time-in-system Variance for a Finite Jobset," *Management Sci.*, 21 (1975), 540–543.

Sundararaghavan, P. and M. Ahmed, "Minimizing the Sum of Absolute Lateness in Single-machine and Multi-machine Scheduling," *Naval Res. Logistics Quarterly*, 31 (1984), 325–333.

Vani, V. and M. Raghavachari, "Deterministic and Random Single Machine Sequencing with Variance Minimization," *Oper. Res.*, 35 (1987), 111–120.

Ventura, J. and X. Weng, "Minimizing the Single-Machine Completion Time Variances," to appear in *Management Sci.*

*Accepted by Hau L. Lee; received December 8, 1993. This paper has been with the authors 7 months for 2 revisions.*