MINIMIZING THE EXPECTED WEIGHTED NUMBER OF TARDY JOBS IN STOCHASTIC FLOW SHOPS

O.J. BOXMA

Centre for Mathematics and Computer Science, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

F.G. FORST

College of Business Administration, Marquette University, Milwaukee, WI 53233, USA

This paper is devoted to two types of stochastic scheduling problems, one involving a single machine and the other involving a flow shop consisting of an arbitrary number of machines. In both problem types, all jobs to be processed have due dates, and the objective is to find a job sequence that minimizes the expected weighted number of tardy jobs. For the single-machine case, sufficient optimality conditions for job sequences are derived for various choices of due date and processing time distributions. For the case of a flow shop with an arbitrary number of machines and identically distributed due dates for all jobs, we prove the following intuitively appealing results: (i) when all jobs have the same processing time distributions, the expected weighted number of tardy jobs is minimized by sequencing the jobs in decreasing order of the weights, (ii) when all weights are equal, the jobs should be sequenced according to an increasing stochastic ordering of the processing time distributions.

stochastic sequencing * tardiness * flow shop

1. Introduction

This paper is devoted to two types of stochastic scheduling problems, one involving a single machine and the other involving a flow shop consisting of an arbitrary number of machines. In both cases the general problem description is as follows. A set of n jobs, $J = (1, 2, \ldots, n)$, has to be processed. All jobs are available at time zero. All jobs have stochastic due dates and deterministic non-negative weights. The objective is to obtain a job sequence $S = (j_1, j_2, \ldots, j_n)$, a permutation of J, that minimizes the expected weighted number of tardy jobs (i.e., jobs that are not completed before their due date). Which conditions (with regard to weights, processing times and due dates) can be formulated that guarantee that S is an optimal job sequence?

Let us first consider the single-machine case in more detail. The processing times X_i of job i, $i = 1, 2, \ldots, n$, are independent, not necessarily identically distributed, stochastic variables. The due dates D_i of job i, $i = 1, 2, \ldots, n$, are independent, not necessarily identically distributed, stochastic variables. The due dates are independent of the processing times. Job i has weight $w_i \geqslant 0$. Let C_i represent the completion time of job i on the machine. When $C_i > D_i$, the decision maker incurs a penalty w_i for job i; otherwise the decision maker incurs no penalty for job i. We define $U_i(S) := 1$ when $C_i > D_i$, and $U_i(S) := 0$ otherwise. In this single-machine case the objective reduces to

$$\min \mathrm{E}[N(S)] := \mathrm{E}\left[\sum_{i=1}^{n} w_i \mathrm{E}(U_i(S))\right]. \tag{1.1}$$

In Section 2 we search for sufficient conditions for S to be an optimal job sequence, for various choices of

distributions of due dates and processing times. For some of these cases we derive simple criteria which completely determine the optimal job sequence, whereas in other cases we are only able to formulate rather complicated and strong sufficient conditions. In the latter cases we also discuss the question whether simple criteria may be expected to hold. In all cases studied, job preëmption is not allowed.

Subsequently we consider the case of a flow shop with an arbitrary number m of machines. Each job of J has to be processed on each machine. The job order is the same for all machines. The processing times $X_{i,j}$ for job i on machine j, $i = 1, \ldots, n$, $j = 1, \ldots, m$, are independent stochastic variables. The due dates D_i of job i, $i = 1, 2, \ldots, n$, are independent, identically distributed stochastic variables. The due dates are independent of the processing times. Job i has weight w_i. Let $C_{i,m}$ represent the completion time of job i on the last machine. When $C_{i,m} > D_i$, the decision maker incurs a penalty w_i for job i; otherwise the decision maker incurs no penalty for job i. We define $U_i(S) := 1$ when $C_{i,m} > D_i$, and $U_i(S) := 0$ otherwise. In this flow-shop case the objective again reduces to (1.1). The usual assumptions for job sequencing problems are made, including non-preëmption of jobs and unlimited storage between machines. For this case we prove the following intuitively appealing results: (i) when all jobs have the same processing time distributions (which may vary from machine to machine), the expected weighted number of tardy jobs is minimized by sequencing the jobs in decreasing order of weights, (ii) when all weights are equal, the jobs should be sequenced according to an increasing stochastic ordering of the processing time distributions.

Notes on related literature

Karp [4] has studied the complexity of the deterministic version of the single-machine tardiness problem. He has shown that the problem of minimizing the weighted number of tardy jobs is binary NP-hard. Recently several results have appeared in the literature for *stochastic* versions of single-machine problems with tardiness.

Balut [1] studied the case in which the job processing times are independent, normally distributed stochastic variables and the objective is to find a job sequence which minimizes the number of tardy jobs (or maximizes the number of early jobs) with probability not smaller than a given constant. The due dates are constant, and the jobs are numbered in ascending due date order. Balut developed an algorithm that he claimed determines the optimal job sequence. In this sequence the early jobs are sequenced in ascending due date order, whereas the tardy jobs are sequenced in any order. However, Kise and Ibaraki [6] presented a counterexample involving three jobs for which Balut's algorithm fails to provide the optimal job sequence. They then proved that the problem Balut formulated is NP-complete, implying that an efficient and exact algorithm for this problem probably does not exist. Kise, Shiomi, Uno and Chao [7] analyzed almost the same problem as Balut, but to avoid an NP-complete problem, they assumed $m_i < m_j$ implies $v_i^2 \leqslant v_j^2$ and $w_i \geqslant w_j$. Here m_i, v_i^2 and w_i are the mean, variance and weight for job i. The objective is to determine the job sequence which minimizes the weighted number of tardy jobs, subject to the constraint that some specified jobs must not be tardy. By modifying Balut's algorithm, Kise et al. developed a straightforward algorithm to determine the optimal job sequence.

Katoh and Ibaraki [5] also considered a problem similar to Balut's, except that all jobs have a common due date d. The objective is to find a job sequence which minimizes the number of tardy jobs with probability not smaller than a given constant which is between 0.5 and 1. The authors developed an algorithm to find the optimal sequence, which can be computed in polynomial time.

Pinedo [8] analysed a single-machine problem in which the job processing times are independent, negative exponentially distributed stochastic variables with rate k_i. Each job has a weight w_i and a stochastic due date; each due date has the same probability distribution (hence the expected due dates are identical for all jobs). Pinedo proves that processing the jobs in decreasing order of $k_i w_i$ minimizes the expected weighted number of tardy jobs. Pinedo's rule is also optimal when the jobs have a common due date d, which may be constant or random.

Glazebrook [3] has derived several results for the single-machine case with due dates, in which job preëmption *is* allowed.

To our knowledge, no results have yet appeared in the literature for the model of Section 3, viz. the flow shop where the objective is to minimize the expected weighted (or unweighted) number of tardy jobs.

2. The single-machine case

Consider the single-machine case described in Section 1. We first discuss a simple special case, to give an indication of the complications that can arise.

Let J consist of $n = 2$ jobs, and let the processing times X_1, X_2 be negative exponentially distributed with mean 1. Let the due dates D_1, D_2 be constants d_1, d_2, respectively. Compare the two sequences $S = (1, 2)$ and $S^* = (2, 1)$ w.r.t. expected weighted number of tardy jobs. Then,

$$\mathrm{E}[N(S)] - \mathrm{E}[N(S^*)]$$

$$= w_1 \Pr\{ X_1 > d_1 \} + w_2 \Pr\{ X_1 + X_2 > d_2 \} - w_1 \Pr\{ X_2 + X_1 > d_1 \} - w_2 \Pr\{ X_2 > d_2 \}$$

$$= -w_1 d_1 \mathrm{e}^{-d_1} + w_2 d_2 \mathrm{e}^{-d_2}. \tag{2.1}$$

To simplify matters further, take $w_1 = w_2$. Then,

$$\mathrm{E}[N(S)] - \mathrm{E}[N(S^*)] = -d_1 \mathrm{e}^{-d_1} + d_2 \mathrm{e}^{-d_2}. \tag{2.2}$$

Job 1 should be processed before job 2 iff the r.h.s. of (2.2) is non-positive. The condition $d_1 \mathrm{e}^{-d_1} \geqslant d_2 \mathrm{e}^{-d_2}$ cannot be further simplified without additional information concerning d_1 and d_2. Indeed, the function $g(z) := z \mathrm{e}^{-z}$ is increasing on $[0, 1)$ and decreasing on $(1, \infty)$; so S is the optimal sequence in the case $d_2 \leqslant d_1 \leqslant 1$, but also in the opposite case $d_2 \geqslant d_1 \geqslant 1$. The first case shows that it is possible that, if two jobs have identical weights and processing time distributions, the job with the smaller due date should be processed last.

Thus warned, let us turn to the general case (different weights, different due date distributions, different processing time distributions). In the following $D_j(\cdot)$ and $F_j(\cdot)$ shall denote the due date distribution and job processing time distribution for job j. We compare $\mathrm{E}[N(S)]$ for job sequence $S = (j_1, \ldots, j_s, i, u, j_{s+3}, \ldots, j_n)$ with $\mathrm{E}[N(S^*)]$ for job sequence $S^* = (j_1, \ldots, j_s, u, i, j_{s+3}, \ldots, j_n)$; so we discuss the effect of interchanging the neighbouring jobs i and u, presently in positions $s + 1$ and $s + 2$. Since all job processing times are independent, it is clear that the expected weighted numbers of tardy jobs in the collections $\{ j_1, \ldots, j_s \}$ and $\{ j_{s+3}, \ldots, j_n \}$ are the same for S and S^*. Therefore we need only consider the expected weighted number of tardy jobs in the collection $\{ i, u \}$. Hence, with C_i^* and C_u^* the completion times of job i and job u, respectively, for S^*,

$$\mathrm{E}[N(S)] - \mathrm{E}[N(S^*)]$$

$$= w_i \Pr\{ C_i > D_i \} + w_u \Pr\{ C_u > D_u \} - w_i \Pr\{ C_i^* > D_i \} - w_u \Pr\{ C_u^* > D_u \}$$

$$= w_i \Pr\{ X_{j_1} + \cdots + X_{j_s} + X_i > D_i \} + w_u \Pr\{ X_{j_1} + \cdots + X_{j_s} + X_i + X_u > D_u \}$$

$$- w_i \Pr\{ X_{j_1} + \cdots + X_{j_s} + X_u + X_i > D_i \} - w_u \Pr\{ X_{j_1} + \cdots + X_{j_s} + X_u > D_u \}. \tag{2.3}$$

Introducing

$$\tilde{F}(t) := F_{j_1}(t) * \cdots * F_{j_s}(t), \qquad t \geqslant 0,$$

we observe that job i should be processed before job u iff

$$\mathrm{E}[N(S^*)] - \mathrm{E}[N(S)] = w_i \int_{t=0}^{\infty} \mathrm{d}D_i(t) [\tilde{F}(t) * F_i(t) - \tilde{F}(t) * F_i(t) * F_u(t)]$$

$$- w_u \int_{t=0}^{\infty} \mathrm{d}D_u(t) [\tilde{F}(t) * F_u(t) - \tilde{F}(t) * F_i(t) * F_u(t)] \geqslant 0. \tag{2.4}$$

Formula (2.4) is the starting-point for our further investigations. Unfortunately, it is too general to allow useful statements. We investigate two important special cases in more detail, viz., I: all due dates are i.i.d. (independent, identically distributed), and II: all job processing times are i.i.d.

I: All due dates are i.i.d.

Assume that $D_i(t) \equiv D(t)$, $i = 1, 2, \ldots, n$. Formula (2.4) reduces to

$$E[N(S^*)] - E[N(S)]$$

$$= \int_{t=0}^{\infty} dD(t)\left[\tilde{F}(t) * \{ w_i F_i(t) - w_i F_i(t) * F_u(t) - w_u F_u(t) + w_u F_i(t) * F_u(t) \}\right] \geq 0. \quad (2.5)$$

We consider four subcases which can be analyzed completely.

Ia: $X_i \sim \exp(\lambda_i)$

Pinedo [8] has studied Case I under the additional assumption that all processing times are negative exponentially distributed with mean $1/\lambda_i$. Formula (2.5) then implies that

$$E[N(S^*)] - E[N(S)] \geq 0 \quad \text{iff} \quad \lambda_i w_i \geq \lambda_u w_u. \quad (2.6)$$

As this condition is transitive, it follows that the jobs have to be sequenced in decreasing order of $\lambda_j w_j$ [8].

Ib: $D_i \sim \exp(d)$

Consider a subcase which in a sense is complementary to Ia: all due dates have the same negative exponential distribution, with mean $1/d$, but the processing times have arbitrary distributions $F_i(\cdot)$ with Laplace–Stieltjes transforms (LST) $f_i(s) := \int_{t=0}^{\infty} e^{-st} \, dF_i(t)$. Formula (2.5) now yields

$$E[N(S^*)] - E[N(S)] \geq 0 \quad \text{iff} \quad w_i f_i(d)(1 - f_u(d)) - w_u f_u(d)(1 - f_i(d)) \geq 0,$$

so

$$E[N(S^*)] - E[N(S)] \geq 0 \quad \text{iff} \quad w_i/[1/f_i(d) - 1] \geq w_u/[1/f_u(d) - 1]. \quad (2.7)$$

As this condition is transitive, we can formulate:

Theorem 2.1. *In the case of i.i.d. due dates with a negative exponential distribution with mean $1/d$, the expected weighted number of tardy jobs is minimized when the jobs are sequenced in decreasing order of $w_j/[1/f_j(d) - 1]$.*

Remark 2.1. The sequencing condition in Theorem 2.1 reduces to the one in (2.6) if the processing times are negative exponentially distributed with mean $1/\lambda_i$.

Ic: $F_i(\cdot) \equiv F(\cdot)$, $i = 1, 2, \ldots, n$

When due dates *and* processing times are i.i.d. (with distributions $D(\cdot)$ and $F(\cdot)$, respectively), then (2.5) reduces to

$$E[N(S^*)] - E[N(S)] = (w_i - w_u)\int_{t=0}^{\infty} dD(t)\left[F^{(s+1)*}(t) - F^{(s+2)*}(t)\right] \geq 0. \quad (2.8)$$

Hence:

Theorem 2.2. *In the case of i.i.d. due dates and i.i.d. processing times, the expected weighted number of tardy jobs is minimized by sequencing the jobs in decreasing order of weights.*

In Subcases Ia, Ib and Ic, simple criteria are obtained which completely determine the optimal job sequence. The following subcase leads to a stronger ordering relation.

Id: $w_i = w$, $i = 1, 2, \ldots, n$

When all weights are identical, (2.5) reduces to

$$E[N(S^*)] - E[N(S)] = w\int_{t=0}^{\infty} dD(t)\left[\tilde{F}(t) * (F_i(t) - F_u(t))\right] \geq 0. \quad (2.9)$$

Introduce the stochastic ordering relation $F_i \leqslant_{st} F_u$, denoting

$$F_i(t) \geqslant F_u(t) \quad \text{for all} \quad t \geqslant 0.$$ (2.10)

See Stoyan [9] for a discussion of this ordering, including some examples. The following theorem immediately follows from (2.9):

Theorem 2.3. *Assume that all due dates are i.i.d. and all weights are identical. If the jobs can be sequenced according to an increasing stochastic ordering, then this yields the minimal expected weighted number of tardy jobs.*

Remark 2.2. Theorem 2.3 is intuitively clear: when all jobs have the same weights and due date distributions, the 'smallest' jobs should be processed first.

The stochastic ordering condition (2.10) is obviously of a different character than Condition (2.7). Contrary to Condition (2.7), Condition (2.10) need not always be fulfilled for (at least) one of the $n!$ possible permutations of J. Unfortunately, Condition (2.10) cannot be weakened to the condition

$$E[X_i] \leqslant E[X_u],$$ (2.11)

as the following trivial example with $n = 2$ jobs shows. Let $w_1 = w_2$ and $D_1 \sim D_2 \sim \exp(1)$; hence both Subcases Ib and Id apply. Let $f_1(s) = \exp(-s)$ and $f_2(s) = 1/[1 + \beta s]$, $s \geqslant 0$: X_1 is constant (equal to one), and X_2 is negative exponentially distributed with mean β. Condition (2.10) is not fulfilled. Application of the simple Condition (2.7) shows that the optimal sequence is $(1, 2)$ iff $E[X_u] = \beta \geqslant e - 1$, which implies that Condition (2.11) cannot be used to determine which of the two orderings is optimal.

Remark 2.3. The foregoing derivations for Case I reveal that, in fact, we do not need the independence of the due dates; it suffices to assume that the stochastic variables $D_1, \ldots, D_n$ are *exchangeable*, i.e., $D_{j_1}, \ldots, D_{j_n}$ have the same joint distribution for all sequences $(j_1, \ldots, j_n)$ (cf. [8]).

II: All job processing times are i.i.d.

Assume that $F_i(t) \equiv F(t)$, $i = 1, 2, \ldots, n$. Formula (2.4) reduces to

$$E[N(S^*)] - E[N(S)] = \int_{t=0}^{\infty} \left[F^{(s+1)*}(t) - F^{(s+2)*}(t) \right] \left[w_i \, \mathrm{d}D_i(t) - w_u \, \mathrm{d}D_u(t) \right] \geqslant 0.$$ (2.12)

One manageable subcase has been discussed in Ic (all due dates i.i.d.). However, as already suggested by the simple example of $n = 2$ jobs in the beginning of this section, it is very hard to make general ordering statements when due dates are *not* identically distributed. We consider two subcases, IIa and IIb, for which some partial results can be obtained.

IIa: Constant due dates; $X_i \sim \exp(\lambda)$

When all due dates are constants, viz., $D_i = d_i$, etc., and all processing times are negative exponentially distributed with mean $1/\lambda$, then (2.12) reduces to (exploit the relation between the negative exponential distribution and the Poisson process)

$$w_i (\lambda d_i)^{s+1} \mathrm{e}^{-\lambda d_i} \geqslant w_u (\lambda d_u)^{s+1} \mathrm{e}^{-\lambda d_u},$$

and, with

$$y_j := \lambda d_j, \quad j = 1, 2, \ldots, n,$$

$$w_i y_i^{s+1} \mathrm{e}^{-y_i} \geqslant w_u y_u^{s+1} \mathrm{e}^{-y_u}$$ (2.13)

(for $s = 0$, compare with (2.1)). This inequality clearly shows the influence of the position of job i (position $s + 1$). Condition (2.13) should hold for neighbouring jobs i and u in *all* possible positions, which leads to

the following two sets of transitive conditions:

$$\begin{cases} y_i \geqslant y_u, \\ w_i y_i e^{-y_i} \geqslant w_u y_u e^{-y_u}, \end{cases} \tag{2.14a}$$

$$\begin{cases} y_i \leqslant y_u, \\ w_i y_i^{n-1} e^{-y_i} \geqslant w_u y_u^{n-1} e^{-y_u}. \end{cases} \tag{2.14b}$$

In other words, don't interchange two neighbouring jobs i and u if (2.14a) holds, and neither if (2.14b) holds; and if (2.14a) holds for all i, u, then it determines an optimal sequence (similarly for (2.14b)).

Note that an optimal sequence need not satisfy (2.14a) or (2.14b); these conditions are sufficient, not necessary. Also note that, in the case of *identical* due dates, (2.14a) and (2.14b) reduce to the criterion $w_i \geqslant w_u$.

IIb: $D_i \sim \exp(d_i)$

Assume that the due date D_i is negative exponentially distributed with mean $1/d_i$, $i = 1, 2, \ldots, n$, and denote the LST of the processing time distribution by $f(\cdot)$. Formula (2.12) reduces to

$$w_i f^{s+1}(d_i)\left(1 - f(d_i)\right) \geqslant w_u f^{s+1}(d_u)\left(1 - f(d_u)\right).$$

Similarly as in IIa we obtain two sets of transitive conditions for job i to be processed before job u,

$$\begin{cases} f(d_i) \geqslant f(d_u), \\ w_i f(d_i)\left(1 - f(d_i)\right) \geqslant w_u f(d_u)\left(1 - f(d_u)\right), \end{cases} \tag{2.15a}$$

$$\begin{cases} f(d_i) \leqslant f(d_u), \\ w_i f^{n-1}(d_i)\left(1 - f(d_i)\right) \geqslant w_u f^{n-1}(d_u)\left(1 - f(d_u)\right). \end{cases} \tag{2.15b}$$

Note that $f(d_i) \geqslant f(d_u)$ iff $d_i \leqslant d_u$.

3. The flow-shop case

We now turn to the case of a flow shop with m machines, as described in Section 1. Each job of a set J of n jobs has to be processed on all machines. The objective again is to minimize $E[N(S)]$, the expected weighted number of tardy jobs. We shall prove flow shop generalizations of results for Subcases Ic and Id of Section 2. First consider a generalization of Ic:

(i) All due dates are i.i.d., the job processing times $X_{i,j}$ of job i on machine j, $i = 1, \ldots, n$, $j = 1, \ldots, m$, are independent, and processing times of different jobs on the same machine are identically distributed.

We shall prove the following generalization of Theorem 2.2:

Theorem 3.1. *In order to minimize the expected weighted number of tardy jobs in case* (i), *the jobs should be sequenced in decreasing order of their weights.*

Proof. We compare $E[N(S)]$ and $E[N(S^*)]$ for the job sequences $S = (j_1, \ldots, j_s, i, u, j_{s+3}, \ldots, j_n)$ and $S^* = (j_1, \ldots, j_s, u, i, j_{s+3}, \ldots, j_n)$; so we discuss the effect of interchanging the neighbouring jobs i and u, presently in positions $s + 1$ and $s + 2$. Since all job processing times are independent, and the processing times of different jobs on the same machine are identically distributed, interchanging jobs i and u will only affect the expected weighted number of tardy jobs in the collection of two jobs $\{i, u\}$. Hence, with $C^*_{i,m}$ and $C^*_{u,m}$ the completion times of job i and job u, respectively, on machine m for the job sequence S^*,

$$E[N(S)] - E[N(S^*)]$$
$$= w_i \Pr\{C_{i,m} > D_i\} + w_u \Pr\{C_{u,m} > D_u\} - w_i \Pr\{C^*_{i,m} > D_i\} - w_u \Pr\{C^*_{u,m} > D_u\}. \tag{3.1}$$

Since all due dates are identically distributed, with distribution $D(\cdot)$,

$$E[N(S)] - E[N(S^*)]$$
$$= \int_{t=0}^{\infty} dD(t)\left[w_i \Pr\{C_{i,m} > t\} + w_u \Pr\{C_{u,m} > t\} - w_i \Pr\{C_{i,m}^* > t\} - w_u \Pr\{C_{u,m}^* > t\}\right].$$

$$(3.2)$$

In order to understand the structure of terms like $\Pr\{C_{i,m} > t\}$, we first consider the case $m = 2$. Then $C_{i,m} = C_{i,2}$ is the maximum of $(s + 1)$ sums of $(s + 2)$ stochastic variables,

$$C_{i,2} = \max\Big[X_{j_1,1} + X_{j_2,1} + \cdots + X_{j_s,1} + X_{i,1} + X_{i,2},$$
$$X_{j_1,1} + X_{j_2,1} + \cdots + X_{j_s,1} + X_{j_s,2} + X_{i,2},$$
$$\cdots$$
$$X_{j_1,1} + X_{j_2,1} + X_{j_2,2} + \cdots + X_{j_s,2} + X_{i,2},$$
$$X_{j_1,1} + X_{j_1,2} + X_{j_2,2} + \cdots + X_{j_s,2} + X_{i,2}\Big],$$

$$(3.3)$$

and similarly,

$$C_{u,2} = \max\Big[X_{j_1,1} + X_{j_2,1} + \cdots + X_{j_s,1} + X_{i,1} + X_{u,1} + X_{u,2},$$
$$X_{j_1,1} + X_{j_2,1} + \cdots + X_{j_s,1} + X_{i,1} + X_{i,2} + X_{u,2},$$
$$\cdots$$
$$X_{j_1,1} + X_{j_2,1} + X_{j_2,2} + \cdots + X_{j_s,2} + X_{i,2} + X_{u,2},$$
$$X_{j_1,1} + X_{j_1,2} + X_{j_2,2} + \cdots + X_{j_s,2} + X_{i,2} + X_{u,2}\Big].$$

$$(3.4)$$

The definition of $C_{i,2}^*$ implies that $C_{i,2}^*$ is obtained from $C_{u,2}$ by interchanging i and u. Hence $C_{i,2}^*$ and $C_{u,2}$ are identically distributed. Similarly, $C_{i,2}$ and $C_{u,2}^*$ are identically distributed. Similar formulas as (3.3) and (3.4) are obtained for general $m \geq 2$. E.g., for $m = 3$ consider $C_{i,3}$ in the $s + 1$ possible cases in which machine 3 is idle for the last time before processing job i; before processing job j_s; ...; before processing job j_1. In the first case, add $X_{i,3}$ to the maximum of all completion time sums corresponding to machines 1 and 2 and jobs $j_1, \ldots, j_s, i$ (the sums in (3.3)); in the second case, add $X_{j_s,3} + X_{i,3}$ to the maximum of all completion time sums corresponding to machines 1 and 2 and jobs $j_1, \ldots, j_s$; etc. Taking the maximum of all these sums amounts to taking the maximum of $(s + 1) + s + (s - 1) + \cdots + 1 = \binom{s+2}{2}$ sums of $(s + 3)$ stochastic variables. We leave it to the reader to check that, for general $m \geq 1$, $C_{i,m}$ can be written as the maximum of $\binom{s+m-1}{m-1}$ sums of $(s + m)$ stochastic variables.

Returning to (3.2) and exploiting the structure of $C_{i,m}$ and its companion terms as sketched above, the following – obvious – results can now be formally proved:

(1) $C_{i,m}$ and $C_{u,m}^*$ are identically distributed,

(2) $C_{u,m}$ and $C_{i,m}^*$ are identically distributed,

(3) $C_{i,m} \leq_{st} C_{u,m}$ (cf. (2.10)), so for all $t \geq 0$, $\Pr\{C_{i,m} > t\} \leq \Pr\{C_{u,m} > t\}$. E.g., from (3.3) and (3.4),

$$C_{u,2} = \max\Big[X_{j_1,1} + X_{j_2,1} + \cdots + X_{j_s,1} + X_{i,1} + X_{u,1} + X_{u,2}, C_{i,2} + X_{u,2}\Big].$$

Finally, from (3.2),

$$E[N(S^*)] - E[N(S)] = (w_i - w_u)\int_{t=0}^{\infty} dD(t)\left[\Pr\{C_{u,m} > t\} - \Pr\{C_{i,m} > t\}\right] \geq 0$$

iff $w_i \geq w_u$.

$$(3.5)$$

This concludes the proof of the theorem. □

(ii) All due dates are i.i.d., $w_i = w$, $i = 1, \ldots, n$, the job processing times $X_{i,j}$ of job i on machine j,

$i = 1, \ldots, n$, $j = 1, \ldots, m$, are independent, and $X_{i,1}, \ldots, X_{i,m}$ are identically distributed with distribution $F_i(\cdot)$, for $i = 1, \ldots, n$ (i.e., a job has the same processing time distribution on each machine).

Analogous to Theorem 2.3 (Subcase Id) we prove:

Theorem 3.2. *If, in case* (ii), *the processing time distribution* $F_k(\cdot)$ *can be sequenced according to an increasing stochastic ordering, this yields the minimal expected weighted number of tardy jobs.*

Proof. Again we compare $E[N(S)]$ and $E[N(S^*)]$ for the job sequences $S = (j_1, \ldots, j_s, i, u, j_{s+3}, \ldots, j_n)$ and $S^* = (j_1, \ldots, j_s, u, i, j_{s+3}, \ldots, j_n)$, discussing the effect of interchanging the neighbouring jobs i and u, presently in positions $s + 1$ and $s + 2$. Although it is clear that interchanging jobs i and u does not affect $C_{j_k,m}$ and $C^*_{j_k,m}$, for $k = 1, \ldots, s$, this is not a priori clear for $k = s + 3, \ldots, n$. We'll return to this point later, for the moment just claiming that (3.2) is indeed still valid. We further claim that $C_{u,m}$ and $C^*_{i,m}$ are identically distributed. For $m = 2$ this readily follows from (3.4): interchanging i and u in (3.4) does not change the last s sums, while the first (second) sum of $C_{u,2}$ has the same distribution as the second (first) sum of $C^*_{i,2}$. For general $v \geqslant 2$ the same result is obtained by carefully exploiting the structure of $C_{u,v}$ and $C^*_{i,v}$ as explained earlier in this section. We thus not only find that $C_{u,m}$ and $C^*_{i,m}$ are identically distributed; we also find that $C_{j_k,m}$ and $C^*_{j_k,m}$, $k \geqslant s$, are identically distributed, which justifies our claim that (3.2) is valid. Note that $C_{j_k,m}$ and $C^*_{j_k,m}$ for $k \geqslant s$ are in general *not* completely identical, contrary to their $k \leqslant s$ counterparts.

It follows from (3.2) that

$$E[N(S^*)] - E[N(S)] = w \int_{t=0}^{\infty} dD(t) [\Pr\{C^*_{u,m} > t\} - \Pr\{C_{i,m} > t\}]. \tag{3.6}$$

$C^*_{u,m}$ is obtained from $C_{i,m}$ by replacing all indices i by u. It trivially follows from $F_i \leqslant_{st} F_u$ that

$$\Pr\{C^*_{u,m} > t\} \geqslant \Pr\{C_{i,m} > t\}, \qquad t \geqslant 0.$$

This concludes the proof of the theorem. $\square$

Remark 3.1. We have no general results for flow shops with non-identically distributed due dates – which is not surprising in view of the far from promising results for the corresponding single-machine models in Section 2. See Forst [2] for an analysis of one particular case with $m = 2$ machines.

References

[1] S.J. Balut, "Scheduling to minimize the number of late jobs when set-up and processing times and uncertain", *Management Science* **19**, 1283–1288 (1973).

[2] F.G. Forst, "Minimizing the expected weighted number of tardy jobs in single-machine and two-machine stochastic flow shops", Report, College of Business Administration, Marquette University, Milwaukee, WI, 1985.

[3] K.D. Glazebrook, "On stochastic scheduling problems with due dates", *Int. J. of Systems Sciences* **14**, 1259–1271 (1983).

[4] R.M. Karp, "Reducibility among combinatorial problems", in: R.E. Miller and J.W. Thatcher, eds., *Complexity of Computer Computations*, Plenum Press, New York, 1972, 85–103.

[5] N. Katoh and T. Ibaraki, "A polynomial time algorithm for a chance-constrained single machine scheduling problem", *Operations Research Letters* **2**, 62–65 (1983).

[6] H. Kise and T. Ibaraki, "On Balut's algorithm and LP-completeness for a chance-constrained scheduling problem", *Management Science* **29**, 384–388 (1983).

[7] H. Kise, A. Shiomi, M. Uno and D. Chao, "An efficient algorithm for a chance-constrained scheduling problem", *J. of the Operations Research Society of Japan* **25**, 193–203 (1982).

[8] M. Pinedo, "Stochastic scheduling with release dates and due dates", *Operations Research* **31**, 559–572 (1983).

[9] D. Stoyan, *Qualitative Eigenschaften und Abschaetzungen Stochastischer Modelle*, Akademie Verlag, Berlin, 1977.00 (1986)