Analytic Approximations for Multiserver Batch-Service Workstations With Multiple Process Recipes in Semiconductor Wafer Fabrication

Ming-Guang Huang, *Member, IEEE*, Pao-Long Chang, and Ying-Chyi Chou

***Abstract*—This study extends previous results for batch-service workstations to batch-service/batch-lot workstations with multiple process recipes, e.g., diffusion operations in semiconductor manufacturing. The model considered herein explicitly considers the existence of a manufacturing operation associated with multiple process recipes in the semiconductor factory. Consequently, the revised balance equations are submitted and an improved approximation is presented for this case. Based on a comparison with simulation results, this new approximation is shown to be superior to the previously developed analytical approaches. This new approximation is especially strong in cases where the number of process recipes grows, system traffic intensity is moderate, and arrival rate of each recipe is nearly the same.**

I. Introduction

THIS STUDY expands upon our previous work [1], which identified four categories of semiconductor wafer fabrication workstation, and presented a detailed analytic process for evaluating system performance measures. The four categories of workstation identified are:

1) single-lot/single-service workstations with single machines ($M/M/1$);
2) single-lot/single-service workstations with multiple machines ($M/M/c$);
3) batch-lot/batch-service workstations ($M/M^x/c$);
4) batch-lot/batch-service workstations with multiple process recipes ($M^y/M^{x,y}/c$).

Notations in parentheses indicate their corresponding queueing model under assumptions of Poisson arrival and exponential service. The upper symbol x, y denotes that the lots are served in batch and with multiple process recipes.

As indicated in [1], the analytic techniques involving the first three categories of workstations are well known and effective. This study focuses on analytic approximation for batch-service/batch-lot workstations with multiple process recipes, aiming to develop a refined version so as to revise and refurbish the corresponding category in [1]. Both this study and [1] aim to replace time-consuming simulation with mathematical planning, while sacrificing a little precision.

Batch-service workstations, which process lots in batches, are an essential type of workstation in semiconductor foundries. Each batch-service workstation has a batch size, which denotes the maximum permissible number of lots that can be grouped into a single batch. Furthermore, batch-service workstations with multiple process recipes exist in semiconductor manufacturing, e.g., in diffusion operations, and possibly involve various processing conditions, such as different processing times, setup operations, and so on. Consequently, a process code is attached to each arriving lot and specifies which lots can be grouped into the same batch for processing. This study follows the rule that the workstation, which may contain multiple servers, begins to process a new batch as soon as a server become available and at least one lot is waiting and groups as many lots as are available in a batch. Additionally, the batch service times are independent and identically exponentially distributed regardless of how many lots the batches contain. Also, the queueing system has infinite buffer capacity to accommodate waiting lots.

Analyzing batch service systems has received considerable attention. Queueing systems with multiserver and batch service have been analyzed in [2]–[5]. Among these studies, Ghare assumed that queueing system are characterized by exponential service time and Poisson arrival. Ghare extended Arora's analytic procedure for the case of two servers to cover the multiserver systems and successfully derived explicit expressions for the equilibrium state queue-length distribution. Ghare [6] later extended his analytic results and finally presented concise expressions for certain performance measures.

Recent batch-service studies are mostly devoted to solve the problems of traffic transportation (see [7]–[10]) and production management (see [11]–[13]). To our knowledge, none of these authors has considered the inevitable problems with various process recipes that occur in semiconductor fabrication systems. Consequently, Connors *et al.* has utilized the analytic results of Ghare and also Cromie and Chaudry for multiserver batch-service workstations (i.e., $M/M^x/c$ queueing system) that approximate performance measures for multiserver batch-service workstations with multiple process recipes (i.e., $M^y/M^{x,y}/c$ queueing system). Connors' suggestion will misestimate the queue-length distribution and, thus, performance measures of $M^y/M^{x,y}/c$ queueing system, particularly when the traffic intensity of the workstation is

Manuscript received June 12, 2000; revised June 11, 2001. This work was supported by the National Science Council, R.O.C., under Grant NSC-89-2622-E-009-002.

M.-G. Huang is with the Department of Finance and Banking, Shih Chien University, Taipei, Taiwan.

P.-L. Chang is with the Institute of Business and Management, National Chiao Tung University, Taipei, Taiwan (e-mail: paolong@cc.nctu.edu.tw).

Y.-C. Chou is with the Department of Business Administration, Ming Chuan University, Taipei, Taiwan.

Publisher Item Identifier S 0894-6507(01)09760-3.

0894–6507/01$10.00 © 2001 IEEE

moderate. Later, Section II uses queue length as an example to further illustrate and verify this comment.

II. Analytic Approach

r	Number of process recipes.
λ_j	Arrival rate of lots for process recipe j.
λ	Total arrival rate: $\lambda = \sum_{j=1}^{r} \lambda_j$.
$E[S_j]$	Batch service time of lots for process recipe j.
$E[S]$	Mean batch service time and service rate $\mu = 1/E[S]$. $E[S]$ can be directly aggregated as follows:

$$E[S] = \sum_{j=1}^{r} \frac{\lambda_j}{\lambda} E[S_j]. \tag{1}$$

μ	Service rate: $\mu = 1/E[S]$.
c	Number of identical batch servers.
w	Maximum batch size.
$p_{m,n}^{(r)}$	Probability that m servers are busy and n lots are queuing with r process recipes.
$z_0^{(r)}$	Root of the derived equation with r process recipes.
$L_q^{(r)}$	Average number of lots in the queue with r process recipes.
$L^{(r)}$	Average number of lots in the system with r process recipes.
$M^{(r)}$	Average number of busy servers with r process recipes.
$\overline{w}^{(r)}$	Average batch size for processing with r process recipes.
$\rho^{(r)}$	Utilization with r process recipes.
$W_q^{(r)}$	Mean queuing time with r process recipes.
$W^{(r)}$	Mean system sojourn time with r process recipes.

A. Analytic Expressions for $M/M^x/c$ Queueing System

First, the queue-length distribution for the $M/M^x/c$ queueing system with $r = 1$ is introduced, based on the analytic results of [4]. This queue-length distribution is derived from the balance equations with $r = 1$, which are reprinted in Appendix A

$$p_{0,0}^{(1)} = \frac{1}{\frac{(\lambda E[S])^c}{c!} \frac{z_0^{(1)}}{z_0^{(1)} - 1} + \sum_{m=0}^{c-1} \frac{(\lambda/\mu)^m}{m!}} \tag{2}$$

$$p_{m,0}^{(1)} = p_{0,0}^{(1)} \frac{(\lambda/\mu)^m}{m!}; \qquad m = 1, 2, \ldots, c-1 \tag{3}$$

$$p_{c,0}^{(1)} = p_{0,0}^{(1)} \frac{(\lambda/\mu)^c}{c!} \tag{4}$$

$$p_{c,n}^{(1)} = p_{c,0}^{(1)} \left(\frac{1}{z_0^{(1)}}\right)^n; \qquad n = 1, 2, \ldots \tag{5}$$

where $z_0^{(1)}$ is a single root, lying in the interval (1, $cw/\lambda E[S]$), of the following derived equation:

$$\frac{\lambda}{c\mu} z^{(1)} - \left(1 + \frac{\lambda}{c\mu}\right) + \left(z^{(1)}\right)^{-w} = 0. \tag{6}$$

Equation (6) is derived from the balance equations. The derived course is omitted herein and readers interested in detail can refer to [4].

Based on the above results, several important performance measures that are primarily derived from the analytic results of [6] and [12]

$$L_q^{(1)} = \sum_{n=0}^{\infty} n p_{c,n} = p_{c,0} \frac{z_0^{(1)}}{\left(z_0^{(1)} - 1\right)^2} \tag{7}$$

$$\begin{aligned} E\left[M^{(1)}\right] &= \sum_{m=0}^{c-1} m p_{m,0} + c \sum_{n=0}^{\infty} p_{c.n} \\ &= \lambda/\mu - c p_{c,0} \left(\frac{z_0^{(1)} \frac{\lambda}{c\mu} - 1}{z_0^{(1)} - 1} \right) \end{aligned} \tag{8}$$

$$\rho^{(1)} = \sum_{m=0}^{c-1} p_{m,0} \frac{m}{c} + \sum_{n=0}^{\infty} p_{c,n} = \frac{E\left[M^{(1)}\right]}{c} \tag{9}$$

$$\overline{w}^{(1)} = \frac{\lambda}{\rho^{(1)} c\mu} = \frac{\lambda}{E\left[M^{(1)}\right]\mu} \tag{10}$$

$$L^{(1)} = L_q^{(1)} + \overline{w}^{(1)} E\left[M^{(1)}\right] = L_q^{(1)} + \lambda/\mu. \tag{11}$$

By using Little's rule, the following measures are obtained:

$$W_q^{(1)} = \frac{L_q^{(1)}}{\lambda} \tag{12}$$

$$W^{(1)} = W_q^{(1)} + E[S]. \tag{13}$$

B. Analytic Approximation for $M^y/M^{x,y}/c$ Queueing System

As mentioned earlier, since the equations in Section II-A do not treat the problem with multiple process recipes explicitly, the errors in queue-length distribution and, thus, performance measures result from using the $M^y/M^x/c$ queueing system to approximate the $M^y/M^{x,y}/c$ queueing system. To provide further evidence for this hypothesis, we estimated queue length in a number of simulation experiments, including $r = 1$, $r = 2$, and $r = 3$, under a variety of total arrival rates. Additionally, all of these experiments were assumed and characterized using the following system parameters:

1) $E[S] = 1/3$ h for $r = 1$, $E[S_1] = 1/2$ h and $E[S_2] = 1/6$ h for $r = 2$, and $E[S_1] = 1/3$ h, $E[S_2] = 1/6$ h, and $E[S_3] = 1/2$ h for $r = 3$.;
2) $c = 3$;
3) $w = 5$;
4) let $\lambda_j = \lambda/r$: $j = 1, \ldots, r$, i.e., the arrival rate of each process recipe is assigned to be identical.

The simulation model assumed both service times and interarrival times (i.e., lots srelease rule) to follow exponential distribution. At any given time, the dispatching rule employed herein selects the process recipe with the most lots for processing. Furthermore, individual process recipes are treated as if they own respective buffers during waiting and grouping into batches.

TABLE I
RESULTS OF QUEUE LENGTH FOR SIMULATION EXPERIMENTS

Total arrival rate (λ)	$z_0^{(1)}$	The number of process recipes				
		$r = 1$	$r = 2$		$r = 3$	
		$L_q^{(1)}$	$L_q^{(2)}$	$\Delta L = (L_q^{(2)} - L_q^{(1)}) / L_q^{(1)}$	$L_q^{(3)}$	$\Delta L = (L_q^{(3)} - L_q^{(1)}) / L_q^{(1)}$
5/hour (0.111*)	2.789	0.124 (0.054**)	0.150 (0.057**)	0.210	0.165 (0.058**)	0.321
7.5/hour (0.167)	2.175	0.362 (0.165)	0.466 (0.173)	0.287	0.567 (0.195)	0.566
10/hour (0.222)	1.860	0.667 (0.303)	0.896 (0.358)	0.343	1.115 (0.527)	0.672
12.5/hour (0.278)	1.663	1.008 (0.402)	1.372 (0.566)	0.361	1.731 (0.745)	0.723
15/hour (0.333)	1.528	1.425 (0.610)	1.918 (0.844)	0.346	2.404 (1.002)	0.709
17.5/hour (0.389)	1.428	1.896 (0.799)	2.494 (1.046)	0.315	3.134 (1.365)	0.663
20/hour (0.444)	1.349	2.422 (1.005)	3.076 (1.332)	0.270	3.896 (1.643)	0.609
22.5/hour (0.5)	1.287	3.075 (1.237)	3.771 (1.676)	0.226	4.582 (2.025)	0.490
25/hour (0.556)	1.234	4.036 (1.717)	4.895 (2.046)	0.213	5.715 (2.456)	0.416
27.5/hour (0.611)	1.191	5.084 (2.143)	5.894 (2.544)	0.179	6.853 (3.062)	0.368
30/hour (0.667)	1.152	6.125 (2.563)	7.114 (3.155)	0.151	8.099 (3.549)	0.322
32.5/hour (0.722)	1.119	7.982 (3.490)	9.047 (4.025)	0.133	9.954 (4.378)	0.247
35/hour (0.778)	1.090	10.678 (4.633)	11.902 (5.252)	0.115	13.003 (5.958)	0.208
37.5/hour (0.833)	1.064	15.085 (6.541)	15.772 (6.984)	0.088	16.318 (7.242)	0.167
40/hour (0.889)	1.041	24.596 (10.097)	25.338 (10.462)	0.065	25.985 (10.730)	0.128

*: traffic intensity, ** : standard deviation.

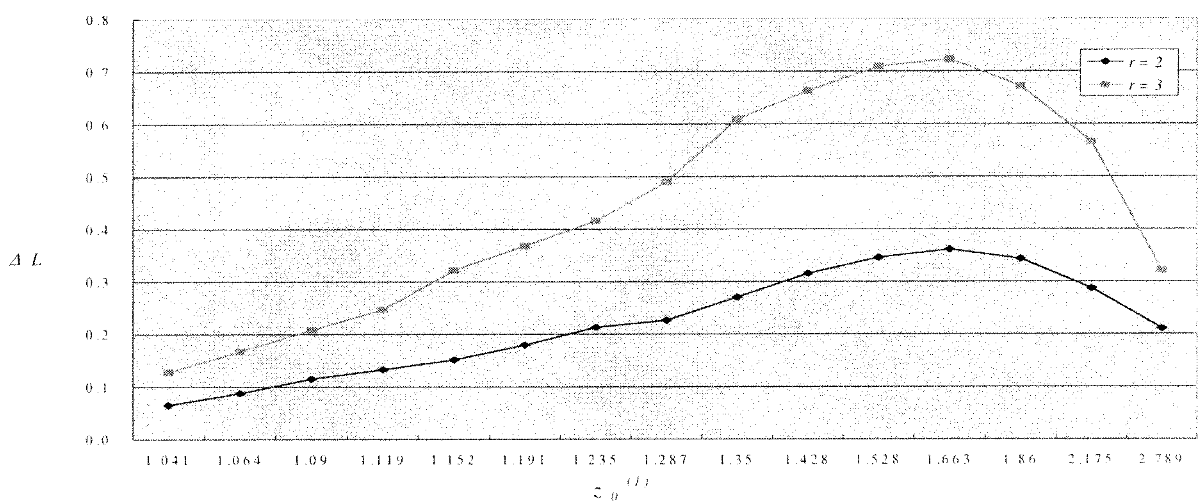

Fig. 1. Relationship between relative increment of queue length and $z_0^{(1)}$.

Table I summarizes the results of the above experiments, where the $z_0^{(1)}$ values are derived from (6) (with only one kind of process recipe being assumed in the system) and ΔL is the relative increment of queue length between $r > 1$ and $r = 1$, given the same total arrival rate. To examine more clearly how queue length varies with traffic intensity and number of process recipes, Fig. 1 illustrates the relationship between marginal increment of queue lengths (ΔL) and values of $z_0^{(1)}$ for these experiments.

Two of the main characteristics of the increasing queue length, apparent from Fig. 1, are:

1) as $z_0^{(1)}$ approaches one (causing traffic intensity to get heavier) or greatly exceeds one (causing traffic intensity to get lighter), queue length only increases slightly;
2) as $z_0^{(1)}$ approaches the middle region, queue length increases accordingly.

These observations are intuitive. The reasons behind these phenomena are briefly explained below and illustrated using an example of an extreme case in which the arrival rates of individual process recipe are all nearly identical.

1) When the arrival rates are very small, servers frequently remain idle. Therefore, the incoming lots are processed almost upon arrival. Consequently, processing batch sizes are around one. In this case, no matter how many kinds of process recipes the incoming lots are classified into, queue length differs little as long provided the total arrival rate remains unchanged.
2) When arrival rates are large, servers will constantly remain busy. Therefore, the arriving lots almost always have to wait for the server and, thus, it will build up a large queue of lots in front of the workstation. Consequently,

TABLE II
COMPARISON OF QUEUE LENGTHS FOR $r = 2$, $\lambda_1 = 0.2\lambda$ AND $\lambda_2 = 0.8\lambda$, AND $E[S_1] = 1$ h AND $E[S_2] = 1/6$ h UNDER SYSTEM CONDITION 1)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
10/hour (0.222*)	0.821 (0.388**)	0.668 (1.8595***)	0.822 (1.7264****)	0.186	-0.001
15/hour (0.333)	1.702 (0.698)	1.449 (1.5280)	1.698 (1.4602)	0.149	0.002
20/hour (0.444)	2.831 (1.217)	2.474 (1.3494)	2.815 (1.3106)	0.126	0.005
25/hour (0.556)	4.407 (1.947)	3.926 (1.2344)	4.388 (1.2111)	0.109	0.004
30/hour (0.667)	6.878 (3.145)	6.260 (1.1524)	6.911 (1.1385)	0.090	-0.005
35/hour (0.778)	11.633 (5.011)	10.829 (1.0901)	11.855 (1.0824)	0.069	-0.019
40/hour (0.889)	25.816 (10.505)	24.388 (1.0406)	26.531 (1.0373)	0.055	-0.028

*: traffic intensity; ** : standard deviation; ***: $z_0^{(I)}$; ****: $z_0^{(r)}$.

since the number of lots for each kind of process recipe in the queue almost exceeds the maximum batch size, all actual batch sizes for processing are approximately equal to the maximum batch size. In this case, the effect resembles 1), i.e., regardless of how many kinds of process recipe the incoming lots are classified into, queue lengths differ little under the consistent total arrival rate.

3) When arrival rates are moderate, server utilization remains normal and queue length is lightly accumulated. The total number of lots in queue likely exceeds the maximum batch size; the number of lots for each kind of process recipe in the queue is often less than the maximum batch size. However, lots cannot be grouped into one batch for processing even if other lots, belonging to different process recipes and possessing different process code, remain in the queue and the processing batch size has not yet reached maximum batch size. In this case, deviation of queue length will be a maximum if $r = 1$ is used to approximate the context of $r > 1$.

Thus, for systems with r ($r > 1$) process recipes, the goal is to determine an adjusted $z_0^{(r)}$, which should reflect the changing pattern of system states under various combinations of number of process recipes and arrival rate of each individual recipe. This goal is achieved by revising Cromie and Chaudry's derived equation, which merely applies to one process recipe. The first two terms of the left-hand side in (6), which imply the arrival component of the incoming and state outgoing, respectively, do not directly relate to the number of process recipes. More specifically, the difference in the number of process recipes only influences the third term of left- hand side in (6). The basis of inference is that (6) assumes that the batch size is equal to w whenever at least one lot remain in the queue after any a service been activated. However, this assumption is only true if one process recipe is considered, while in systems with more than one process recipe, the batch size might be between one and w when the number of lots remaining in the queue is between one and $(r-1)(w-1)$. Consequently, we consider replacing (6) with

$$\frac{\lambda}{c\mu} z^{(r)} - \left(1 + \frac{\lambda}{c\mu}\right) + \sum_{k=1}^{w} e_k \left(z^{(r)}\right)^{-k} = 0 \qquad (14)$$

where $e_k (k = 1, 2, \ldots, w)$ denotes the proportion of the service batch size that equals k if at least one lot is waiting in the queue.

Meanwhile, for computing $e_k (k = 1, 2, \ldots, w)$, it is helpful to understand the balance equations that contain r ($r > 1$)process recipes, as outlined in Appendix B. Based on the balance equations, the computation of $e_k (k = 1, 2, \ldots, w)$ is now presented in Appendix C.

After incorporating the computation of $e_k (k = 1, 2, \ldots, w)$ into (14) to numerically solve $z_0^{(r)}$, the queue-length distribution can be obtained as can performance measures with r process recipes via the same computational processes as in (2)–(13), with $z_0^{(r)}$ instead of $z_0^{(1)}$.

III. NUMERICAL EXAMPLES

This section attempts to examine the scale of error by using the proposed analytic process. Accordingly, numerical experiments were undertaken for $r = 2$ and 3. Because of considerations of space, only the results of queue length obtained from this new approximation and from the analytical approach presented by Cromie and Chaudhry [6] are compared with the simulation results.

These experiments were classified into two categories according to differences in system conditions, as follows.

1) The number of identical batch servers: $c = 3$; maximum batch size: $w = 5$ lots.
2) The number of identical batch servers: $c = 5$; maximum batch size: $w = 3$ lots.

The experiments were then conducted on two levels according to discrepancies in the arrival rates of individual process recipes.

1) *Large differences in arrival rates:* Queue length for $r = 2$ and 3 is compared in Tables II–V, respectively.
2) *Small difference between arrival rates:* Comparisons of queue length for $r = 2$ and 3 are tabulated in Tables VI–IX, respectively.

Tables II–IX clearly show that the proposed analytic expressions are more accurate than that suggested by Cromie and Chaudry, especially when the number of process recipes rises, traffic intensities are moderate, and arrival rates of process recipes remain nearly the same. From these comparisons, although the results tend to be slightly overestimated and the

TABLE III
COMPARISON OF QUEUE LENGTHS FOR $r = 2$, $\lambda_1 = 0.2\lambda$ AND $\lambda_2 = 0.8\lambda$, AND $E[S_1] = 1$ h AND $E[S_2] = 1/6$ h UNDER SYSTEM CONDITION 2)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
15/hour (0.333*)	0.640 (0.304**)	0.555 (1.8393***)	0.634 (1.7579****)	0.133	0.009
20/hour (0.444)	1.356 (0.625)	1.200 (1.5477)	1.345 (1.4992)	0.115	0.008
25/hour (0.556)	2.407 (1.096)	2.166 (1.3631)	2.392 (1.3332)	0.100	0.006
30/hour (0.667)	4.098 (1.648)	3.731 (1.2338)	4.079 (1.2156)	0.089	0.005
35/hour (0.778)	7.229 (3.113)	6.791 (1.1370)	7.366 (1.1269)	0.061	-0.019
40/hour (0.889)	16.700 (7.560)	15.844 (1.0613)	17.087 (1.0569)	0.051	-0.023
42/hour (0.933)	29.037 (11.881)	27.864 (1.0353)	29.990 (1.0328)	0.040	-0.033

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

TABLE IV
COMPARISON OF QUEUE LENGTHS FOR $r = 3$, $\lambda_1 = 0.1\lambda$, $\lambda_2 = 0.2\lambda$, AND $\lambda_3 = 0.7\lambda$, AND $E[S_1] = 1/3$ h, $E[S_2] = 5/6$ h, AND $E[S_3] = 1/6$ h UNDER SYSTEM CONDITION 1)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
10/hour (0.222*)	0.977 (0.427**)	0.668 (1.8595***)	0.982 (1.6290****)	0.316	-0.005
15/hour (0.333)	2.023 (0.887)	1.449 (1.5280)	2.008 (1.3977)	0.284	0.008
20/hour (0.444)	3.254 (1.225)	2.474 (1.3494)	3.287 (1.2696)	0.240	-0.010
25/hour (0.556)	4.962 (2.018)	3.926 (1.2344)	5.067 (1.1842)	0.209	-0.021
30/hour (0.667)	7.638 (3.192)	6.260 (1.1524)	7.914 (1.1215)	0.180	-0.036
35/hour (0.778)	12.913 (5.665)	10.829 (1.0901)	13.480 (1.0727)	0.161	-0.044
40/hour (0.889)	28.521 (11.762)	24.388 (1.0406)	29.987 (1.0331)	0.145	-0.051

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

TABLE V
COMPARISON OF QUEUE LENGTHS FOR $r = 3$, $\lambda_1 = 0.1\lambda$, $\lambda_2 = 0.2\lambda$, AND $\lambda_3 = 0.7\lambda$, AND $E[S_1] = 1/3$ h, $E[S_2] = 5/6$ h, AND $E[S_3] = 1/6$ h UNDER SYSTEM CONDITION 2)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
15/hour (0.333*)	0.735 (0.327**)	0.555 (1.8393***)	0.744 (1.6697****)	0.246	-0.010
20/hour (0.444)	1.548 (0.718)	1.200 (1.5477)	1.562 (1.4414)	0.225	-0.009
25/hour (0.556)	2.692 (0.964)	2.166 (1.3631)	2.747 (1.2953)	0.195	-0.020
30/hour (0.667)	4.481 (1.704)	3.731 (1.2338)	4.641 (1.1916)	0.167	-0.036
35/hour (0.778)	7.956 (3.287)	6.791 (1.1370)	8.318 (1.1131)	0.147	-0.045
40/hour (0.889)	18.218 (8.005)	15.844 (1.0613)	19.173 (1.0509)	0.130	-0.052
42/hour (0.933)	31.592 (12.569)	27.864 (1.0353)	33.587 (1.0294)	0.118	-0.065

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

deviations grow up with traffic intensities, the relative errors of the proposed approximation were found to remain within 12% for both $r = 2$ and $r = 3$. It is fair to say that the approximations still give acceptable precision even if the traffic intensity increases.

As for other performance measures, such as waiting time, utilization etc., since these measures are highly correlated with queue length through $z_0^{(r)}$ and Little's formula, they should have roughly similar error scales.

IV. CONCLUSION

Multiserver batch-service workstations with multiple process recipes, e.g., diffusion operations, are an essential type of workstation in semiconductor manufacturing. An analytic approach for such workstations was developed to evaluate semiconductor fabrication performance. The analytic approximation of queue-length distribution and performance measures of the multiserver batch-service workstations with multiple process recipes is pre-

TABLE VI
COMPARISON OF QUEUE LENGTHS FOR $r = 2$, $\lambda_1 = 0.4\lambda$ AND $\lambda_2 = 0.6\lambda$, AND $E[S_1] = 3.5/6$ h AND $E[S_2] = 1/6$ h UNDER SYSTEM CONDITION 1)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
10/hour (0.222*)	0.878 (0.400**)	0.668 (1.8595***)	0.876 (1.6904****)	0.239	0.002
15/hour (0.333)	1.875 (0.892)	1.449 (1.5280)	1.862 (1.4248)	0.227	0.007
20/hour (0.444)	3.119 (1.225)	2.474 (1.3494)	3.125 (1.2824)	0.207	-0.002
25/hour (0.556)	4.766 (1.833)	3.926 (1.2344)	4.898 (1.1902)	0.176	-0.027
30/hour (0.667)	7.453 (3.187)	6.260 (1.1524)	7.739 (1.1242)	0.160	-0.038
35/hour (0.778)	12.648 (5.524)	10.829 (1.0901)	13.296 (1.0737)	0.144	-0.051
40/hour (0.889)	27.897 (11.399)	24.388 (1.0406)	29.779 (1.0333)	0.126	-0.067

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

TABLE VII
COMPARISON OF QUEUE LENGTHS FOR $r = 2$, $\lambda_1 = 0.4\lambda$ AND $\lambda_2 = 0.6\lambda$, AND $E[S_1] = 3.5/6$ h AND $E[S_2] = 1/6$ h UNDER SYSTEM CONDITION 2)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
15/hour (0.333*)	0.692 (0.316**)	0.555 (1.8393***)	0.690 (1.7080****)	0.198	0.003
20/hour (0.444)	1.476 (0.692)	1.200 (1.5477)	1.464 (1.4654)	0.187	0.008
25/hour (0.556)	2.603 (1.002)	2.166 (1.3631)	2.599 (1.3100)	0.168	0.002
30/hour (0.667)	4.328 (1.546)	3.731 (1.2338)	4.419 (1.2004)	0.138	-0.021
35/hour (0.778)	7.691 (3.044)	6.791 (1.1370)	7.959 (1.1179)	0.117	-0.035
40/hour (0.889)	17.676 (7.638)	15.844 (1.0613)	18.415 (1.0529)	0.103	-0.042
42/hour (0.933)	30.708 (12.655)	27.864 (1.0353)	32.292 (1.0305)	0.092	-0.051

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

TABLE VIII
COMPARISON OF QUEUE LENGTHS FOR $r = 3$, $\lambda_1 = 0.2\lambda$, $\lambda_2 = 0.3\lambda$, AND $\lambda_3 = 0.5\lambda$, AND $E[S_1] = 1/6$ h, $E[S_2] = 1/6$ h, AND $E[S_3] = 1/2$ h UNDER SYSTEM CONDITION 1)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
10/hour (0.222*)	1.086 (0.511**)	0.668 (1.8595***)	1.125 (1.5632****)	0.385	-0.036
15/hour (0.333)	2.295 (0.857)	1.449 (1.5280)	2.377 (1.3429)	0.369	-0.035
20/hour (0.444)	3.793 (1.397)	2.474 (1.3494)	3.952 (1.2275)	0.348	-0.042
25/hour (0.556)	5.829 (2.366)	3.926 (1.2344)	6.147 (1.1534)	0.326	-0.054
30/hour (0.667)	9.054 (3.872)	6.260 (1.1524)	9.657 (1.1002)	0.308	-0.067
35/hour (0.778)	15.191 (6.745)	10.829 (1.0901)	16.517 (1.0595)	0.287	-0.087
40/hour (0.889)	33.277 (13.238)	24.388 (1.0406)	36.859 (1.0269)	0.267	-0.107

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

sented herein. From numerical experiments, the level of estimation error proposed herein is below 12% for systems with two and three process recipes. While the proposed approximation is characterized by slight overestimation, acceptable precision is still obtained. In practice, the number of process recipes rarely exceeds three.

Additionally, a simple and effective way to extend the analytic results of $M^y/M^{x,y}/c$ queueing model to the approximations of general $GI^y/G^{x,y}/c$ model is to incorporate the correction factor suggested by [14] when the first two moments (i.e., mean and variance) of interarrival times of aggregated lots and batch service times are known. The computation of correction factor depends on size of υ^a (the squared coefficient of variation of interarrival times for aggregated lots) and υ^s (the squared coefficient of variation of batch service times). Due to space consideration, readers who are interested in such expansion and transformation can refer to [1] and [14] for details.

Finally, discussing how the novel analytical approximation developed herein is actually applied in semiconductor manufacturing is also worthwhile. To this end, the relevant

TABLE IX
COMPARISON OF QUEUE LENGTHS FOR $r = 3$, $\lambda_1 = 0.2\lambda$, $\lambda_2 = 0.3\lambda$, AND $\lambda_3 = 0.5\lambda$, AND $E[S_1] = 1/6$ h, $E[S_2] = 1/6$ h, AND $E[S_3] = 1/2$ h UNDER SYSTEM CONDITION 2)

Total arrival rate (λ)	Queue length			Relative errors	
	Simulation	Analytical results			
		Cromie and Chandhry	Huang et al.	Cromie and Chandhry	Huang et al.
15/hour (0.333*)	0.821 (0.377**)	0.555 (1.8393***)	0.855 (1.6019****)	0.324	-0.041
20/hour (0.444)	1.762 (0.842)	1.200 (1.5477)	1.806 (1.3915)	0.320	-0.025
25/hour (0.556)	3.068 (1.036)	2.166 (1.3631)	3.179 (1.2599)	0.294	-0.036
30/hour (0.667)	5.075 (1.973)	3.731 (1.2338)	5.364 (1.1678)	0.265	-0.057
35/hour (0.778)	8.948 (3.774)	6.791 (1.1370)	9.598 (1.0987)	0.241	-0.073
40/hour (0.889)	20.180 (8.292)	15.844 (1.0613)	22.093 (1.0443)	0.215	-0.094
42/hour (0.933)	34.692 (14.041)	27.864 (1.0353)	38.665 (1.0256)	0.197	-0.115

*: traffic intensity; ** : standard deviation; ***: $z_0^{(l)}$; ****: $z_0^{(r)}$.

manufacturing data, such as the number of process recipes, arrival rate of each process recipe, service time of each process recipe, maximal batch size, number of servers, and so on, must be gathered in advance to serve as inputs of queueing model. The above is done where the arrival rate of each process recipe can be extracted from information including product mix, process routings, and lot/wafer release rules by using a set of translating and aggregating processes. These processes have been described and implemented in [1], so readers interested in further details of these processes can refer to this study. Furthermore, we reiterate that batch-service workstations with multiple process recipes are only one of the types of workstation that is commonplace in the semiconductor manufacturing environment. While workstations can be considered in isolation if necessary, they are more frequently considered as part of whole system and act as part of a whole system being explored.

APPENDIX A
BALANCE EQUATIONS WITH OR WITHOUT ONE PROCESS RECIPE [4]

The balance equations with or without one process recipe are given as follows:

$$\begin{aligned}(d/dt)P_{c,n}^{(1)}(t) &= -(\lambda + c\mu)P_{c,n}^{(1)}(t) + \lambda P_{c,n-1}(t) \\ &\quad + c\mu P_{c,n+w}^{(1)}(t) \qquad (n > 0) \\ (d/dt)P_{c,0}^{(1)}(t) &= -(\lambda + c\mu)P_{c,0}^{(1)}(t) + \lambda P_{c-1,0}^{(1)}(t) \\ &\quad + c\mu \sum_{k=1}^{w} P_{c,k}^{(1)}(t) \\ (d/dt)P_{m,0}^{(1)}(t) &= -(\lambda + m\mu)P_{m,0}^{(1)}(t) + \lambda P_{m-1,0}^{(1)}(t) \\ &\quad + (m+1)\mu P_{m+1,0}^{(1)}(t) \qquad (1 \le m < c) \\ (d/dt)P_{0,0}^{(1)}(t) &= -\lambda P_{0,0}^{(1)}(t) + \mu P_{1,0}^{(1)}(t). \end{aligned} \tag{15}$$

$$\begin{aligned}
&(d/dt)p_{c,n(l_1,\ldots,l_r)}^{(r)}(t) = -(\lambda + c\mu)p_{c,n(l_1,\ldots,l_r)}^{(r)}(t) + \lambda p_{c,n-1(l_1,\ldots,l_j-1,\ldots,l_r)}^{(r)}(t) + c\mu p_{c,n+w(l_1,\ldots,l_j+w,\ldots,l_r)}^{(r)}(t) \\
&\qquad n > (r-1)(w-1) \\
&\begin{cases}
(d/dt)p_{c,n(l_1,\ldots,l_r)}^{(r)}(t) = -(\lambda + c\mu)p_{c,n(l_1,\ldots,l_r)}^{(r)}(t) + \lambda p_{c,n-1(l_1,\ldots,l_j-1,\ldots,l_r)}^{(r)}(t) + c\mu p_{c,n+w(l_1,\ldots,l_j+w,\ldots,l_r)}^{(r)}(t) \\
\qquad 1 \le n \le (r-1)(w-1); \quad l_j > 0, \quad j = 1, 2, \ldots, r \\
(d/dt)p_{c,n(l_1,\ldots,l_r)}^{(r)}(t) = -(\lambda + c\mu)p_{c,n(l_1,\ldots,l_r)}^{(r)}(t) + \lambda p_{c,n-1(l_1,\ldots,l_j-1,\ldots,l_r)}^{(r)}(t) + c\mu \times \\
\left(\sum_{j=1}^{r} \left[\underset{\text{if } 0<l_j\le n \text{ and } l_j+w \ge \max_{s\in\{1,2,\ldots,r\}, s\ne j}\{l_s\}}{p_{c,n+w(l_1,\ldots,l_j+w,\ldots,l_r)}^{(r)}(t)} + \sum_{k=\max_{s\in\{1,2,\ldots,r\}, s\ne j}\{l_s\}}^{w} \underset{\text{if } l_j=0}{p_{c,n+k(l_1,\ldots,l_{j-1},k,l_{j+1},\ldots,l_r)}^{(r)}(t)} \right] \right) \\
\qquad 1 \le n \le (r-1)(w-1); \quad \text{at least one of } l_j = 0, \quad j = 1, 2, \ldots, r
\end{cases} \\
&(d/dt)p_{c,0(0,\ldots,0)}^{(r)}(t) = -(\lambda + c\mu)p_{c,0(0,\ldots,0)}^{(r)}(t) + \lambda p_{c-1,0(0,\ldots,0)}^{(r)}(t) + c\mu \sum_{j=1}^{r}\sum_{k=1}^{w} p_{c,k(0,\ldots,l_{j-1}=0,k,l_{j+1}=0,\ldots,0)}^{(r)}(t) \\
&(d/dt)p_{m,0(0,\ldots,0)}^{(r)}(t) = -(\lambda + m\mu)p_{m,0(0,\ldots,0)}^{(r)}(t) + \lambda p_{m-1,0(0,\ldots,0)}^{(r)}(t) + (m+1)\mu p_{m+1,0(0,\ldots,0)}^{(r)}(t) \\
&\qquad 1 \le m < c \\
&(d/dt)p_{0,0(0,\ldots,0)}^{(r)}(t) = -\lambda p_{0,0(0,\ldots,0)}^{(r)}(t) + \mu p_{1,0(0,\ldots,0)}^{(r)}(t).
\end{aligned} \tag{16}$$

APPENDIX B
BALANCE EQUATIONS WITH r ($r > 1$) PROCESS RECIPES

The balance equations with r ($r > 1$) process recipes are given in (16) at the bottom of p. 401, where $P^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)}(t)$ denotes that, at time t, c servers (i.e., all servers) are busy and n lots are queuing in which process recipe 1 has l_1 lots, process recipe 2 has l_2 lots, etc., and $\sum_{j=1}^{r} l_j = n$ with r process recipes. As the simulation model stated earlier, the above equations assume that the system will select the process recipe with the greatest number of lots among all process recipes into processing at any given point of time.

APPENDIX C
COMPUTATION OF e_k ($k = 1, 2, \ldots, w$)

At first, $P^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)}(t)$ may be approximated by

$$P^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)}(t) = \frac{n!}{l_1!l_2!\cdots l_r!}\prod_{j=1}^{r}\left(\frac{\lambda_j}{\lambda}\right)^{l_j} P^{(r)}_{c,\,n}(t). \tag{17}$$

Taking the limit as $t \to \infty$, let $P^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)}$ and $P^{(r)}_{c,\,n}$ denote steady-state probabilities of $P^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)}(t)$ and $P^{(r)}_{c,\,n}(t)$, re-

---

$$\alpha_n \cong \sum_{\text{at least one of } l_j=0,\ j=1,2,\ldots,r} p^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)}, \qquad \text{for } n = 1, 2, \ldots, (r-1)(w-1), \tag{19}$$

$$\begin{aligned}\alpha_{n,\,k} \cong &\sum_{\text{at least one of } l_j=0,\ j=1,2,\ldots,r} p^{(r)}_{c,\,n(l_1,\,\ldots,\,l_r)} \\ &\times \sum_{j=1}^{r}\left[\frac{\underset{\text{if } l_j=0}{p^{(r)}_{c,\,n+k(l_1,\,\ldots,\,l_{j-1},\,k,\,l_{j+1},\,\ldots,\,l_r)}}}{\underset{\text{if } l_j>0 \text{ and } l_j+w\ge \max_{s\in\{1,2\ldots,r\},\,s\ne j}\{l_s\}}{p^{(r)}_{c,\,n+w(l_1,\,\ldots,\,l_j+w,\,\ldots,\,l_r)}} + \sum_{\hat{k}=\max_{s\in\{1,2\ldots,r\},\,s\ne j}\{l_s\}}^{w} \underset{\text{if } l_j=0}{p^{(r)}_{c,\,n+\hat{k}(l_1,\,\ldots,\,l_{j-1},\,\hat{k},\,l_{j+1},\,\ldots,\,l_r)}}}\right], \\ &k = 1, 2, \ldots, w-1\end{aligned} \tag{20}$$

$$\begin{aligned}\alpha_{n,\,w} \cong &\sum_{\text{at least one of } l_j=0,\ j=1,2,\ldots,r} p^{(r)}_{c,\,n(l_1,\,l_2,\,\ldots,\,l_r)} \\ &\times \sum_{j=1}^{r}\left[\frac{\underset{\text{if } l_j>0 \text{ and } l_j+w\ge \max_{s\in\{1,2\ldots,r\},\,s\ne j}\{l_s\}}{p^{(r)}_{c,\,n+w(l_1,\,\ldots,\,l_j+w,\,\ldots,\,l_r)}} + \underset{\text{if } l_j=0}{p^{(r)}_{c,\,n+w(l_1,\,\ldots,\,l_{j-1},\,w,\,l_{j+1},\,\ldots,\,l_r)}}}{\underset{\text{if } l_j>0 \text{ and } l_j+w\ge \max_{s\in\{1,2\ldots,r\},\,s\ne j}\{l_s\}}{p^{(r)}_{c,\,n+w(l_1,\,\ldots,\,l_j+w,\,\ldots,\,l_r)}} + \sum_{\hat{k}=\max_{s\in\{1,2\ldots,r\},\,s\ne j}\{l_s\}}^{w} \underset{\text{if } l_j=0}{p^{(r)}_{c,\,n+\hat{k}(l_1,\,\ldots,\,l_{j-1},\,\hat{k},\,l_{j+1},\,\ldots,\,l_r)}}}\right], \\ &\text{for} \quad n = 1, 2, \ldots, (r-1)(w-1)\end{aligned} \tag{21}$$

and

$$\begin{cases} \beta_w \cong \displaystyle\sum_{n=0}^{\infty} p^{(r)}_{c,\,n} - p^{(r)}_{c,\,0} - \sum_{j=1}^{(r-1)(w-1)} \alpha_j + \sum_{j=1}^{(r-1)(w-1)} \alpha_{j,\,w} \\ \beta_k \cong \displaystyle\sum_{j=1}^{(r-1)(w-1)} \alpha_{j,\,k}, \qquad k = 1, 2, \ldots, w-1 \end{cases} \tag{22}$$

$$\begin{aligned}\alpha_n \cong &\,p^{(r)}_{c,\,0}\left(\sum_{l_1=0}^{n}\sum_{l_2=0}^{n-l_1}\cdots\sum_{l_{r-1}=0}^{n-\sum_{u=1}^{r-2} l_u}\left\{\underset{\text{at least one of } l_j=0,\ j=1,2,\ldots,r}{\frac{n!}{l_1!\cdots\left(n-\sum_{u=1}^{r-1} l_u\right)!}\left[\prod_{j=1}^{r-1}\left(\frac{\lambda_j}{\lambda}\right)^{l_j}\left(\frac{\lambda_r}{\lambda}\right)^{(n-\sum_{u=1}^{r-1} l_u)}\right]}\right\}\frac{1}{\left(z_0^{(r)}\right)^n}\right) \\ =&\,p^{(r)}_{c,\,0}\hat{\alpha}_n; \qquad 1 \le n \le (r-1)(w-1)\end{aligned} \tag{23}$$

$$
\begin{aligned}
\alpha_{n,k} \cong\, & p_{c,0}^{(r)} \underbrace{\sum_{l_1=0}^{n} \sum_{l_2=0}^{n-l_1} \cdots \sum_{l_{r-1}=0}^{n-\sum_{u=1}^{r-2} l_u}}_{\text{at least one of } l_j = 0,\, j=1,2,\ldots,r} \left\{ \frac{n!}{l_1!\cdots\left(n-\sum_{u=1}^{r-1} l_u\right)!} \left[\prod_{j=1}^{r-1} \left(\frac{\lambda_j}{\lambda}\right)^{l_j} \left(\frac{\lambda_r}{\lambda}\right)^{(n-\sum_{u=1}^{r-1} l_u)} \right] \frac{1}{\left(z_0^{(r)}\right)^n} \right\} \\
& \times \sum_{j=1}^{r} \left[\frac{\begin{array}{c} \dfrac{(n+k)!}{l_1!\cdots l_{j-1}!k!l_{j+1}!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{k} \right] \dfrac{1}{\left(z_0^{(r)}\right)^{n+k}} \\ \text{if} \quad l_j = 0 \end{array}}{\begin{array}{c} \dfrac{(n+w)!}{l_1!\cdots(l_j+w)!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{l_j+w} \right] \dfrac{1}{\left(z_0^{(r)}\right)^{n+w}} \\ \text{if} \quad l_j > 0 \quad \text{and} \quad l_j + w \geq \displaystyle\max_{\substack{s\in\{1,2,\ldots,r\}\\ s\neq j}} \{l_s\} \\ + \displaystyle\sum_{\hat{k}=\max_{\substack{s\in\{1,2,\ldots,r\}\\ s\neq j}}\{l_s\}}^{w} \left[\dfrac{(n+\hat{k})!}{l_1!\cdots l_{j-1}!\hat{k}!l_{j+1}!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{\hat{k}} \right] \dfrac{1}{\left(z_0^{(r)}\right)^{n+\hat{k}}} \right] \\ \text{if} \quad l_j = 0 \end{array}} \right] \\
\cong\, & p_{c,0}^{(r)} \hat{\alpha}_{n,k}; \qquad k = 1, 2, \ldots, w-1
\end{aligned}
\tag{24}
$$

$$
\begin{aligned}
\alpha_{n,w} \cong\, & p_{c,0}^{(r)} \underbrace{\sum_{l_1=0}^{n} \sum_{l_2=0}^{n-l_1} \cdots \sum_{l_{r-1}=0}^{n-\sum_{u=1}^{r-2} l_u}}_{\text{at least one of } l_j = 0, \quad j=1,2,\ldots,r} \left\{ \frac{n!}{l_1!\cdots\left(n-\sum_{u=1}^{r-1} l_u\right)!} \left[\prod_{j=1}^{r-1} \left(\frac{\lambda_j}{\lambda}\right)^{l_j} \left(\frac{\lambda_r}{\lambda}\right)^{(n-\sum_{u=1}^{r-1} l_u)} \right] \frac{1}{\left(z_0^{(r)}\right)^n} \right\} \\
& \times \sum_{j=1}^{r} \left[\frac{\begin{array}{c} \dfrac{(n+w)!}{l_1!\cdots(l_j+w)!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{l_j+w} \right] \dfrac{1}{\left(x_0^{(r)}\right)^{n+w}} \\ \text{if} \quad l_j > 0 \quad \text{and} \quad l_j + w \geq \displaystyle\max_{\substack{s\in\{1,2,\ldots,r\}\\ s\neq j}} \{l_s\} \\ + \dfrac{(n+w)!}{l_1!\cdots l_{j-1}!w!l_{j+1}!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{w} \right] \dfrac{1}{\left(x_0^{(r)}\right)^{n+w}} \\ \text{if} \quad l_j = 0 \end{array}}{\begin{array}{c} \dfrac{(n+w)!}{l_1!\cdots(l_k+w)!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{l_j+w} \right] \dfrac{1}{\left(z_0^{(r)}\right)^{n+w}} \\ \text{if} \quad l_j > 0 \quad \text{and} \quad l_j + w \geq \displaystyle\max_{\substack{s\in\{1,2,\ldots,r\}\\ s\neq j}} \{l_s\} \\ + \displaystyle\sum_{\hat{k}=\max_{\substack{s\in\{1,2,\ldots,r\}\\ s\neq j}}\{l_s\}}^{w} \left[\dfrac{(n+\hat{k})!}{l_1!\cdots l_{j-1}!\hat{k}!l_{j+1}!\cdots l_r!} \left[\displaystyle\prod_{\substack{s=1\\ s\neq j}}^{r} \left(\frac{\lambda_s}{\lambda}\right)^{l_s} \left(\frac{\lambda_j}{\lambda}\right)^{\hat{k}} \right] \dfrac{1}{\left(z_0^{(r)}\right)^{n+\hat{k}}} \right] \\ \text{if} \quad l_j = 0 \end{array}} \right] \\
\cong\, & p_{c,0}^{(r)} \hat{\alpha}_{n,w},
\end{aligned}
\tag{25}
$$

$$\text{for } n = 1, 2, \ldots, (r-1)(w-1); \quad \text{and}$$

$$\begin{aligned} \beta_w &\cong \sum_{n=0}^{\infty} p_{c,n}^{(r)} - p_{c,0}^{(r)} - \sum_{j=1}^{(r-1)(w-1)} \alpha_j + \sum_{j=1}^{(r-1)(w-1)} \alpha_{j,w} \\ &\cong p_{c,0}^{(r)} \left(\frac{z_0^{(r)}}{z_0^{(r)} - 1} - 1 - \sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_k + \sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_{j,w} \right), \end{aligned} \tag{26}$$

$$\beta_k \cong \sum_{j=1}^{(r-1)(w-1)} \alpha_{j,k} = p_{c,0}^{(r)} \left(\sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_{j,k} \right); \qquad k = 1, 2, \ldots, w-1 \tag{27}$$

$$e_k = \frac{\beta_k}{\sum_{\hat{k}=1}^{w} \beta_{\hat{k}}} = \frac{\sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_{j,k}}{\left(\frac{z_0^{(r)}}{z_0^{(r)} - 1} - 1 - \sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_j + \sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_{j,k} \right) + \sum_{j=1}^{(r-1)(w-1)} \hat{\alpha}_{j,w}}, \quad k = 1, 2, \ldots, w \tag{28}$$

spectively. Meanwhile, the steady-state queue-length distribution is given as

$$P_{c,n}^{(r)} = P_{c,0}^{(r)} \left(\frac{1}{z_0^{(r)}} \right)^n. \tag{18}$$

Let α_n denote the probability that after any service is activated, on average, n $(1 \leq n \leq (r-1)(w-1))$ lots remained in the queue and at least one of $l_j (j = 1, \ldots, r)$ is zero. Let $\alpha_{n,k}$ denote the probability that after any service has been activated, on average, n $(1 \leq n \leq (r-1)(w-1))$ lots remain in the queue, at least one of $l_j (j = 1, \ldots, r)$ is zero, and the actual batch size equals k. Also, let β_k denote the probability that after any a service is activated, on average, the actual batch size equals k. According to the balance equations in (16), these parameters can be approximately obtained as shown in (19)–(22) at the bottom of p. 402. Consequently, the following can be found in (23) at the bottom of p. 402, (24) and (25) on p. 403, and (26) and (27) on p. 404.

Finally, the following approximation is obtained in (28), as shown at the top of the page.

ACKNOWLEDGMENT

The authors are grateful to United Microelectronics Corporation for offering manufacturing process information of semiconductor fabrication.

REFERENCES

[1] P. L. Chang, M. G. Huang, and Y. C. Chou, "Evaluating system performances for semiconductor fabrication using open general queueing networks," *Int. J. Oper. Quant. Manage.*, vol. 4, no. 3, pp. 327–342, Dec. 1998.

[2] K. L. Arora, "Two-server bulk-service queueing process," *Oper. Res.*, vol. 12, pp. 286–295, 1964.

[3] P. B. Rose, "A many server bulk queue," *Oper. Res.*, vol. 14, pp. 1037–1044, 1966.

[4] P. M. Ghare, "Multichannel queueing system with bulk service," *Oper. Res.*, vol. 16, pp. 189–192, 1968.

[5] R. F. Love, "Steady-state solution of the queueing system $E_w/M/S$ with batch service," *Oper. Res.*, vol. 18, pp. 185–189, 1970.

[6] M. V. Cromie and M. L. Chaudhry, "Analytically explicit results for the queueing system $M/M^x/c$ with charts and tables for certain measures of efficiency," *Oper. Res. Quart.*, vol. 27, no. 3, pp. 733–745, 1989.

[7] A. S. Alfa, "Time-inhomogeneous bulk server queue in discrete time: A transportation problem," *Oper. Res.*, vol. 30, no. 4, pp. 650–658, Aug. 1982.

[8] M. T. Juan and D. Worthington, "Practical methods for queue length behavior for bulk service queues of the form $M/G^{0.C}/1$ and $M(t)/G^{0.C}/1$," *Eur. J. Oper. Res.*, vol. 74, pp. 103–113, 1994.

[9] W. B. Powell, "Iterative algorithm for bulk arrival, bulk service queues with Poisson and Non-Poisson arrivals," *Transport. Sci.*, vol. 20, pp. 65–80, 1986.

[10] W. B. Powell and P. Humblet, "The bulk service queue with a general control strategy: Theoretical analysis and a new computational procedure," *Oper. Res.*, vol. 34, pp. 267–275, Mar.–Apr. 1986.

[11] G. R. Bitran and D. Tirupati, "Approximations for product departures from a single-server station with batch processing in multi-product queues," *Manage. Sci.*, vol. 35, no. 7, pp. 851–878, July 1989.

[12] D. P. Connors, G. E. Feigin, and D. D. Yao, "A queueing network model for semiconductor manufacturing," *IEEE Tran. Semiconduct. r Manufact.*, vol. 9, pp. 412–427, Aug. 1996.

[13] P. P. Wang, "Markovian queueing models with periodic review," *Comput. Oper. Res.*, vol. 23, no. 8, pp. 741–754, 1996.

[14] W. Whitt, "Approximations for the $GI/G/m$ Queue," *Prod. Oper. Manage.*, vol. 2, no. 2, pp. 114–161, Spring 1993.

[15] D. Gross and C. M. Harris, *Fundmentals of Queueing Theory*. New York: Wiley, 1985.

Ming-Guang Huang (M'96) received the B.S. degree in computer science and information engineering, the M.S. degree in industrial engineering, and the Ph.D. in business and management from the National Chiao Tung University, Taipei, Taiwan, in 1991, 1992, and 2001, respectively.

He is an Associate Professor with the Department of Finance and Banking, Shih Chien University, Taipei, Taiwan. His current research interests include production planning and management, manufacturing system evaluation, and information management.

Pao-Long Chang received the B.S. degree in mathematics from Fu-Jen Catholic University, Taipei, Taiwan, in 1971, the M.A. degree in mathematics from the State University of New York, Albany, in 1975, and the Ph.D. degree in mathematics from the University of Washington, Seattle, in 1980.

He is a Professor with the Institute of Business and Management, National Chiao Tung University, Taiwan. His previous articles have appeared in the *Journal of the Operations Research Society*, *Japan*, *Journal of the Operational Research Society*, *Journal of Environmental Management*, *Computers and Operations Research*, *International Journal of Production Economics*, *Technovation*, *International Journal of Technology Management and* IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT. His current research interests are in the areas of operations research and technology management.

Ying-Chyi Chou received the B.S. degree in mathematics from the National Tsing Hua University, Hsinchu, Taiwan, in 1994, the M.S. degree in statistics from the National Chung Cheng University, Chiayi, Taiwan, in 1996, and the Ph.D. degree in business and management from National Chiao Tung University, Taipei, Taiwan, in 2001

She is an Assistant Professor with the Department of Business Administration, Ming Chuan University, Taipei, Taiwan. Her current research interests are in the areas of statistics, production planning and management, and inventory management.