Modeling and Analysis of Re-entrant Production Systems

Young Jae Jang and Stanley B. Gershwin
Massachusetts Institute of Technology

***Abstract*— This paper presents a model and analysis of a re-entrant production line with finite buffers and unreliable machines. Semiconductor device and liquid crystal display (LCD) fabrication processes are characterized as a re-entrant process, in which a similar sequence of processing step is repeated several times. The purpose of this paper is to present mathematical formulations and algorithms to analyze the material behavior of the re-entrant production system using the decomposition method. In developing equations for the two-machine building blocks for the re-entrant production line, we modify the existing decomposition model that has been created for the multiple-part type line.**

I. Introduction

This paper presents a mathematical model and analysis of a re-entrant production system that consists of unreliable machines and finite buffers located between machines. Typical examples of such a re-entrant production system are semiconductor device chip and liquid crystal display (LCD) panel fabrication systems whose processes involves a large number of steps with significant number of re-entrant flow paths. In the re-entrant production system, material visits to particular machines or groups of machines several times before it leaves the system. This re-entrant flow behavior with the stochastic nature of the system caused by machine failure or demand changes makes the system difficult predict and analyze.

Gershwin[1] introduced a decomposition method that analyzes the behavior of the manufacturing system with a stochastic queuing model. This method models a manufacturing system as a flow line with unreliable machines and finite buffers. Since then several different variations of decomposition methods have been introduced. A production line with Assembly/Disassembly was investigated with the decomposition method by [2] and a line with loop system using the decomposition method was introduced by [3] and [5]. A production system processing Multiple-part type was introduced by Jang and Gershwin[4]. However, all the decomposition methods developed until now were based on the assumption that parts that was processed by a machine once never re-entered to the machine again. There has been no model considering the re-entrant flow behavior constructed so far.

In this paper we develop an analytical model for a production line with re-entrant flow by applying the decomposition method. As a first step in modeling the re-entrant line, we restrict ourselves to the case that parts re-enters to the system only once. It is found that the material flow behavior of the re-entrant production line is very similar with that of the multiple-part type line. Therefore, we apply the model already developed for the multiple-part type line by Jang and Gershwin[4] to the decomposition model for the re-entrant production line. Therefore, instead of developing a model from scratch, we utilize the equations that we developed in [4] and modify them to represent the re-entrant flow behavior. In this regard, the modeling of re-entrant systems is an extension of the two-part type production system done in [4] and we follow the notations and assumptions we made in [4].

The next section introduces the notations and assumptions of the model we develop. Section III presents the decomposition method for the re-entrant production line. In this section, we compare the flow behaviors between the two-part type line and re-entrant flow line and explain how we apply the decomposition method developed for the two-part type line to the re-entrant flow line. Section IV discuss the derivations of equations for the re-entrant flow, which is the critical equations that makes us possible to utilize the two-part type line concept for the re-entrant production line. Section V presents the numerical results and discuss the quantitative behavior of the system.

II. Modeling

A. Notations and assumptions

Figure 1 represents a re-entrant production line. The line consists of two kinds of components: processing machines M_i denoted by the squares and finite-capacity storage buffers $B_{i,j}$ for work in process inventory, denoted by the circles. Let us define K to be the number of machines. At the beginning and end of the line, there are supply machines, M_0 and demand machines, M_K. Once parts are entered to the line through the supply machine, they are first processed by the machine from M_1 to M_K. During this processing steps, parts are stored in the buffer $B_{i,2}$, $i \in \{0...K\}$. Then they are processed by the *re-entrant machine*, $M_{K+1,2}$. The function of this machine is sending parts back to the machine M_1 so that parts can go through the same processing steps again. This re-entering machine can be either an actual processing or an imaginary machine that logically creates the re-entrant loop. We call the processing steps before the re-entrant machine *Stage 1* ($S = 1$), while the processing steps after the re-entrant machine *Stage 2* ($S = 2$). In this paper we strict out model to $S \in \{1, 2\}$. During the second stage of the process, parts are stored in $B_{i,1}$, $i \in \{1...K\}$.

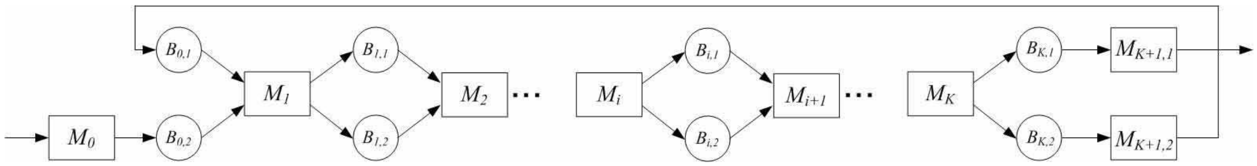

Fig. 1. Re-entrance production line model

Machines, M_i, $i \in \{1...K\}$ are switching between stage 1 and stage 2 processed. We assume that there is no set-up time incurred when the machines switch processing stage. We assume that all the machines in the line are unreliable. Let α denote the state of a machine. If $\alpha = 1$, the machine is said to be *up* or *working*. If $\alpha = 0$, the machine is said to be *down* or *failed*. The state variable representing the state of the machine at the end of time t is written $\alpha_i(t)$. We make the assumption that all the machines in the line have *homogeneous processing times*. That is, the lengths of time that parts spend in machines are fixed, known in advance, and the same for all the machines. For convenience, the processing times are assumed to be scaled to unity. Furthermore, we assume that the yield of all machines is 100%. That is, we do not allow the scrapping or rework of parts.

We assume that all buffers have finite size. The size of buffer $B_{i,j}$ is denoted $N_{i,j}$, where i indicates the production sequence, and $j = 1$ or 2, represents the production stage. We denote the current level of $B_{i,j}$ at the end of time t by $n_{i,j}(t)$. Therefore, $0 \leq n_{i,j}(t) \leq N_{i,j}$, for all (i, j), and for all $t \geq 0$. We make the assumptions that the supply machine is never starved and the demand machines is never blocked.

B. Part Priority Policy

Since each machine in the production line must choose which stage of part to work on when it has a choice, we are required to state a policy by which that choice is made. Our assumption is that each machine will work on stage 2 parts whenever the machine is up, the upstream buffer for stage 2 parts is not empty, and the downstream buffer for stage 2 parts is not full. Each machine will only work on stage 1 parts if it is up, and either blocked or starved for stage 2 parts, and not starved or blocked for stage 1 parts. Under this priority rule, we can possibly achieve a low inventory level by pushing out the parts spent longer time in the system.

C. Machine Parameters and Dynamics

As mentioned earlier, all machines in the line are assumed to be unreliable. We further assume that machines cannot fail if they are idle. This is called *operation dependent failures*. It means that a machine cannot fail if it is either starved or blocked for parts.

All machines are assumed to have geometrically distributed up and down times. We assume that the probability that M_i fails is the same, regardless of the stage of parts the processing machine is working on. We let r_i represent the probability that M_i is up in time $t+1$, given it was down in time t. Likewise, p_i represents the probability that M_i is down in time $t + 1$, given it was up and not blocked or starved in time t. For M_i, the machine parameters can be written as:

$$\begin{aligned} r_i &= Pr\left[\alpha_i(t+1) = 1 | \alpha_i(t) = 0\right] \\ p_i &= Pr\left[\alpha_{i,1}(t+1) = 0 | \right. \qquad (1) \\ &\quad \{\alpha_{i,1}(t) = 1 \cap n_{i-1,1}(t) > 0 \cap n_{i,1}(t) < N_{i,1}\} \cup \\ &\quad \{\alpha_{i,1}(t) = 1 \cap (n_{i-1,1}(t) = 0 \cup n_{i,1}(t) = N_{i,1}) \\ &\quad \left. \cap n_{i-1,2}(t) > 0 \cap n_{i,2}(t) < N_{i,2}\}\right] \\ &\quad \text{for } i = 1, \ldots, K \end{aligned}$$

Likewise, for the supply, demand, and re-entrant machines, the machine parameters are defined as:

$$\begin{aligned} r_0 &= Pr\left[\alpha_0(t+1) = 1 | \alpha_0(t) = 0\right] \\ p_0 &= Pr\left[\alpha_0(t+1) = 0 | \alpha_0(t) = 1 \cap n_{0,2}(t) < N_{0,2}\right] \end{aligned}$$

$$\begin{aligned} r_{K+1,1} &= Pr\left[\alpha_{K+1,1}(t+1) = 1 | \alpha_{K+1,1}(t) = 0\right] \qquad (2) \\ p_{K+1,1} &= Pr[\alpha_{K+1,1}(t+1) = 0 | \\ &\qquad \alpha_{K+1,1}(t) = 1 \cap n_{K,1}(t) > 0] \end{aligned}$$

$$\begin{aligned} r_{K+1,2} &= Pr\left[\alpha_{K+1,2}(t+1) = 1 | \alpha_{K+1.2}(t) = 0\right] \\ p_{K+1,2} &= Pr[\alpha_{K+1,2}(t+1) = 0 | \\ &\qquad \alpha_{K+1.2}(t) = 1 \cap n_{K,2}(t) > 0 \cap n_{0,1} < N_{0,1}] \end{aligned}$$

D. Performance measures

We consider two performance measures in analyzing the re-entrant production line: production rate (throughput rate) and average buffer level.

III. Decomposition Method

A. General idea of the decomposition

We use the decomposition method to analyze the behavior of the re-entrant production line. The decomposition method breaks down the larger system into analytically tractable two-machine one-buffer lines called *two-machine building blocks* or simply *building blocks* and capture the local behavior of the original line, as seen by an observer in a buffer, by

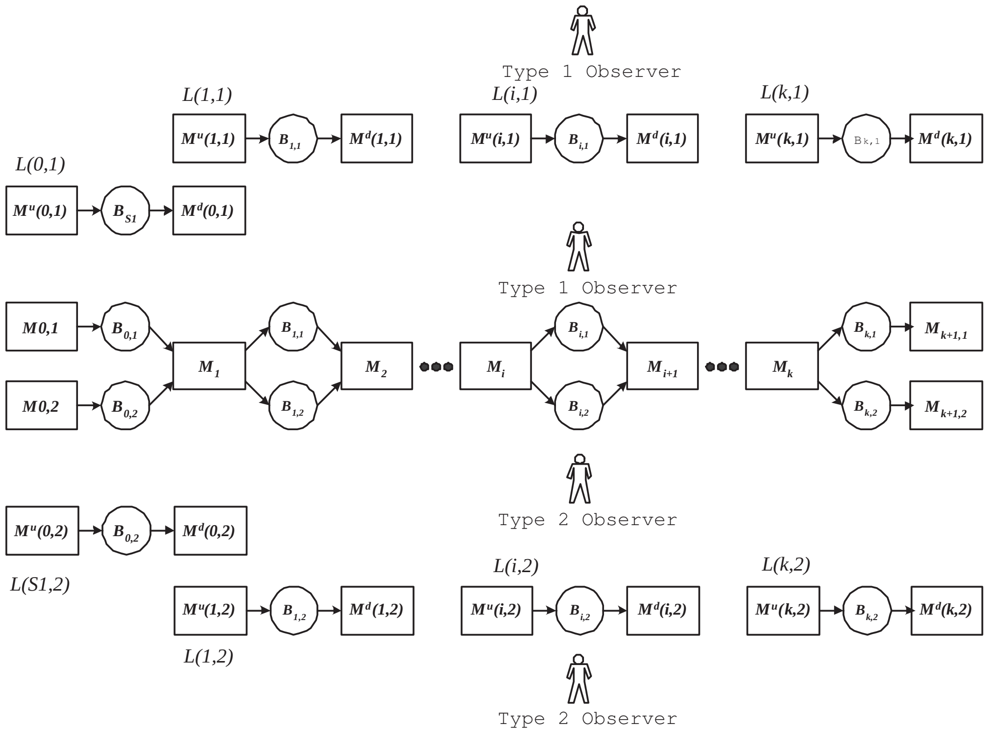

Fig. 2. The decomposition of a line into two-machine lines

choosing appropriate parameters of the two-machine building blocks. Note that each of the two-machine building blocks is constructed with a buffer that is the same size as that of one of the buffers in the original line. The equations that relates the flow behaviors between the original line and two-machine building blocks are called *decomposition equations*.

Figure 2 shows the re-entrant production system decomposed into several two-machine building blocks. As shown in the figure, the inflow and outflow behavior of material an observer in buffer $B_{i,j}$ could see is modeled by the two-machine building blovk, $L(i,j)$.

Note that there are two different types of observers in the figure: stage 2 observers and Stage 1 observers. The Stage 2 observers watch the inflow and outflow of the stage 2 parts while the stage 1 observers watch those of the stage 1 parts. Therefore, we have two different types of building blocks. In the figure, the two-machine building blocks in the top imitate the flow behavior exclusively for the second stage production flow, while the bottom building blocks imitate the flow behavior exclusively for the first stage production flow. For example, let us consider the case that M_i is up, and $n_{i-1,2} = 0$, $n_{i+1,2} < N$, $n_{i-1,1} > 0$, and $n_{i,1} < N_{i.1}$ — second stage part is not available but the first stage part is. Due to the priority rule, the machine will first try to work on the second stage part but it will find that there is no part available in $n_{i-1,2}$. Therefore it will eventually work on the next priority part, which is the first stage part. From the observers' view points on this situation, the observer in $B_{i,1}$ will believe that her upstream machine is down since she does not see any material coming into $B_{i,2}$. On the contrary, the observer in $B_{i,2}$ will believe that her upstream machine is up since this observer sees the part coming into the buffer $B_{i,1}$.

B. Two-part type line vs. re-entrant line decomposition

Before we move onto the detailed modeling of the decomposition equations let us consider the decomposition of the two-part type production line studied by [4]. Figure 2 represents a production line processing two different part types. Machines $M_{0,1}$ and $M_{K+1,1}$ process only Type 1 parts, while machines $M_{0,2}$ and $M_{K+1,2}$ process only Type 2 parts. Each machine, other than the supply and demand machines, process both art types. We assume that there is no set-up time incurred when the machines switch production from one part type to another. When M_i completes work on a part, it sends the part to a buffer downstream of the machine. Each part type has a distinct buffer after each machine. Therefore, a Type 1 part processed at M_i would be sent to $B_{i,1}$. A Type 2 part processed at the same machine would be sent to $B_{i,2}$.

In the two-part type line, since each machine in the production line must choose which part to work on when it has a choice, we are required to state a policy by which that choice is made. Our assumption is that each machine will work on Type 1 parts whenever the machine is up, the upstream buffer for Type 1 parts is not empty, and the downstream buffer for Type 1 parts is not full. Each machine will only work on Type 2 parts if it is up, and either blocked or starved for Type 2 parts, and not starved or blocked for Type 2 parts. Since there are two independent part types in the line, we need to evaluate the production rate for Type 1 and Type 2, that is E_1 and E_2, respectively.

If we examine the flow behavior of the two-part type production line, we can find that there are a lot of similarities between two production lines. First, both lines consist of unreliable machines and finite buffers and also they operate under strict priority rules. Only difference is the presence of the re-entrant line and number of supply and demand machines.

Then, we may ask the following question: *is there any way we can take advantage of this similarity in deriving equations for the re-entrant line instead of constructing equations from scratch?* Here is one approach we propose. Suppose that in the two-part type production line, the parameters for the demand machine for Type 2, $M_{k+1,2}$, are given such that the machine imitates the flow behavior of Type 1 part in the line. Also, at the same time, the parameters for the supply machine for Type 1, $M_{0,1}$, are assigned such that the machine imitates the flow behavior of Type 2 part in the line. In this case, Type 1 and Type 2 flow will imitate the flow behavior for Stage 1 part and Stage 2 part, respectively. Since the Stage 1 and Stage 2 parts are a physically single product type, the following equality is made due to the conservations of the flow:

$$E_1 = E_2 \tag{3}$$

If we apply the above approach to the decomposition method for the re-entrant line, all the decomposition equations are the same as those constructed in two-part type line except the decomposition equations for $L(k,2)$ and $L(0,1)$. In the two-part type line, the downstream machine parameters for $L(k,2)$ are the same as the demand machine for Type 2. Likewise, the upstream machine parameters for $L(0,1)$ are those of the supply machine for Type 1. Notice that the parameters for the actual machines in the line, including supply and demand machines, are independent variables to the system, and therefore, we do not need equations for the downstream machine of $L(k,2)$ and the upstream machine of $L(0,1)$. However, for the re-entrant production line, the parameters for these machines are not independent variables anymore. Therefore, we need to construct a set of equations to match the flow behavior between these two part types. From now on, Type 1 part refers to the Stage 2 part, while Type 2 part refers to the Stage 1 part.

IV. Decomposition Equations for Re-entrant Flow

As mentioned in the previous section, the model we construct for re-entrant system is an extension of the two-part type system model we introduced in [4]. We only need to derive the equations for $L(0,1)$ and $L(k,2)$. We follow the same notations described in [4] for two-machine building blocks. The following list summarizes the building block notations:

- $M^u(i,j)$: Upstream machine in (i,j)
- $M^d(i,j)$: Downstream machine in $L(i,j)$

1) Interruption of flow: For the interruption of flow for $M^u(0,1)$, we use the balance equation:

$$\sum_{i=1}^{3} Ps_i(K,2) r_i^u(K,2) = W^d(K-1,2) p_*^d \tag{4}$$

,where p_*^d is the probability that $M_{K+1,2}$ becomes starved due to any machine failure upstream of $B_{K,2}$. Then

$$\begin{aligned} p_1^u(0,1) &= p_*^d + p_{K+1,2} \qquad (5)\\ &= \frac{1}{W^d(1,2)} \sum_{i=1}^{3} Ps_i(K,2) r_i^u(K,2) + p_{K+1,2} \end{aligned}$$

Similarly, for $M^d(2,2)$

$$\sum_{i=1}^{3} Pb_i(0,1) r_i^d(0,1) = W^u(1,1) p_*^u \tag{6}$$

$$\begin{aligned} p_1^d(K,2) &= p_*^u + p_{K+1,2} \qquad (7)\\ &= \frac{1}{W^u(1,1)} \sum_{i=1}^{3} Pb_i(0,1) r_i^d(0,1) + p_{K+1,2} \end{aligned}$$

2) Resumption of flow: Flow rate idle time is used for the derivations of the resumption of flow equations.

$$E = e_{K+1,2}\left(1 - \widetilde{Ps}(K,2) - \widetilde{Pb}(0,1)\right) \tag{8}$$

where $e_{K+1,2} = \frac{r_{K+1,2}}{r_{K+1,2}+p_{K+1,2}}$, $\widetilde{Ps} = \sum_{i=1}^{3} Ps_i$ and $\widetilde{Pb} = \sum_{i=1}^{3} Pb_i$. Also we know that

$$E^u(0,1) = e^u(0,1)\left(1 - \widetilde{Pb}(0,1)\right)$$

$$E^d(K,2) = e^d(K,2)\left(1 - \widetilde{Ps}(K,2)\right)$$

These can be written

$$\widetilde{Ps}(K,2) = 1 - \frac{E^d(K,2)}{e^d(K,2)}$$

$$\widetilde{Pb}(0,1) = 1 - \frac{E^u(0,1)}{e^u(0,1)}$$

Then (8) becomes

$$E = e_{K+1,2}\left(\frac{E^d(K,2)}{e^d(K,2)} + \frac{E^u(0,1)}{e^u(0,1)} - 1\right)$$

or since $E = E^d(K,2) = E^u(1,1)$,

$$1 = e_{K+1,2}\left(\frac{1}{e^d(K,2)} + \frac{1}{e^u(0,1)} - \frac{1}{E}\right) \tag{9}$$

We know that

$$\frac{1}{e^u(0,1)} = \frac{p^u(0,1) + r^u(0,1)}{r^u(0,1)}$$

$$\frac{1}{e^d(K,2)} = \frac{p^d(K,2) + r^d(K,2)}{r^d(K,2)}$$

Then (9) becomes

$$1 = e_{K+1,2}\left(\frac{p^u(0,1) + r^u(0,1)}{r^u(0,1)} + \frac{p^d(K,2) + r^d(K,2)}{r^d(K,2)} - \frac{1}{E}\right)$$

$$1 = e_{K+1,2}\left(\frac{p^u(0,1)}{r^u(0,1)} + \frac{p^d(K,2)}{r^d(K,2)} - \frac{1}{E} + 2\right)$$

That is,

$$\frac{1}{e_{K+1,2}} + \frac{1}{E} - 2 = \frac{p^u(0,1)}{r^u(0,1)} + \frac{p^d(K,2)}{r^d(K,2)} \tag{10}$$

Two equation are introduced

$$I^u(0,1) = \frac{p^u(0,1)}{r^u(0,1)} \quad \text{and} \quad I^d(K,2) = \frac{p^d(K,2)}{r^d(K,2)} \tag{11}$$

Then we can rewrite (10) such that,

$$\begin{aligned} I^u(0,1) &= \frac{1}{E^d(K,2)} + \frac{1}{e_{K+1,2}} - I^d(K,2) - 2 \\ I^d(K,2) &= \frac{1}{E^u(0,1)} + \frac{1}{e_{K+1,2}} - I^u(0,1) - 2 \end{aligned}$$

A. Algorithm

The algorithm is based on DDX algorithm which is first introduced by [1]. In the re-entrant production line case, we first sweep the high priority line, calculating the upstream two-machine parameter for $M^u(1,1)$, using the parameters of the previous low priority line, and then sweep the low priority line to calculate the downstream two-machine line, $M^d(K-1,2)$ parameters. We then repeat the process for each successive part type.

V. Numerical Results

In order to verify the analytic equation derived in the previous section, we compare the numerical results of a small system with four machines and four buffers with simulations. The small system is shown in Figure 3.

Two separate cases are presented. For the both cases, the machine parameter of M_4 varies, while the rest of machine parameters remains constant. We examine the response of the production rate of the system to the varying parameter and compare the results with simulations.

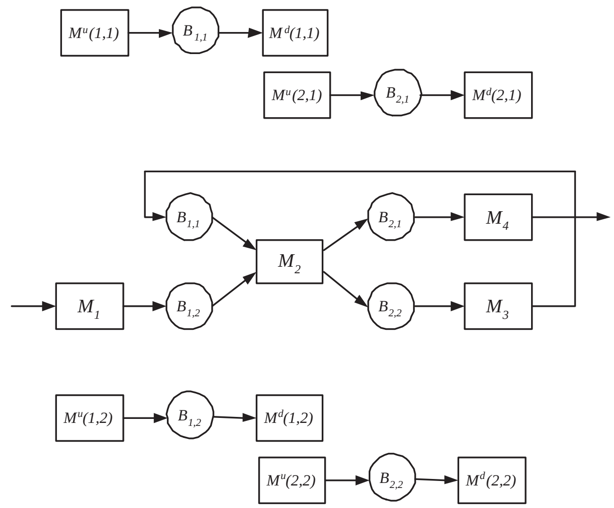

Fig. 3. Simple re-entrance production line model

A. Case1: Varying p_4 and r_4 (e_4 constant)

The system parameter is shown in the Table V-A. For this case, we increase the failure rate of M_4 from 0.3 to 0.52. Also, we vary the repair rate of M_4 to satisfy the isolated production rate of M_4 remains 0.48. The rest of the parameters are unchanged. The result of this case is shown in Figure 4. In the figure, the straight line represents the the production rate of the analytical result and the star and circle marks represent the upper and lower bounds of 95% confidence intervals evaluated from simulation runs. As shown in the figure, the production rate of the system is little bit blow of 0.45. This result matches with our expectation, because although the parameters of M_4 are changed, the isolated production rate of the machine remained the same. Also the bottleneck machine of the system is M_2 and therefore the parameter change of the non-bottleneck machine M_4 has little influence the production rate of the system.

From the figure, we can see that the analytical results are within the upper and lower bounds of the 95% confidence intervals. We calculated the percent error of the production rate from the simulated production rate in the following manner.

$$\%Error = 100 \times \frac{E_{analytical} - E_{sim}}{E_{sim}}$$

The result is shown in Figure 5. As shown in the figure, the most of errors are within 1.5% and the maximum error is about % 2.5.

B. Case2: Varying p_4 with changing e_4

The system parameter of the second case is shown in Table V-A. In this case, we vary p_4 from 0.1 to 0.8. However, unlike the first case, we fix the value r_4, therefore, the isolated production rate for M_4 decreases as p_4 increases. The result is of the case is shown in Figure 6. The production rate

		Case1		Case2	
Machine	Parameter	Value	Iso. Prod. Rate	Value	Iso. Prod. Rate
M_1	r_1	0.48	0.48	0.48	0.48
	p_1	0.52		0.52	
M_2	r_2	0.1	0.9091/2	0.1	0.9091/2
	p_2	0.11	= 0.4545	0.11	=0.4545
M_3	r_3	0.48	0.48	0.48	0.48
	p_3	0.52		0.52	
M_4	r_3	varying	0.48	0.48	varying
	p_3	0.3∼0.52		0.1∼0.8	

TABLE I

CASE1 AND CASE2 PARAMETERS. ($N_1 = N_2 = N_3 = N_4 = 20$)

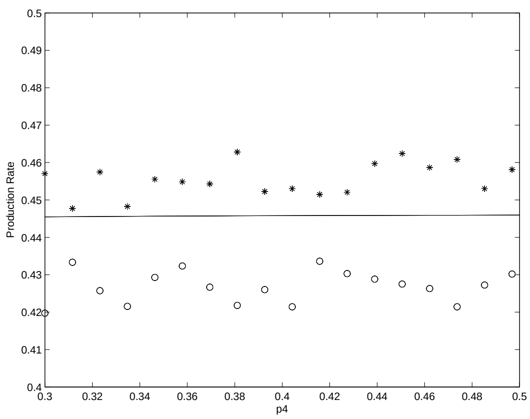

Fig. 4. Production rate vs. p_4 (e_4 fixed)

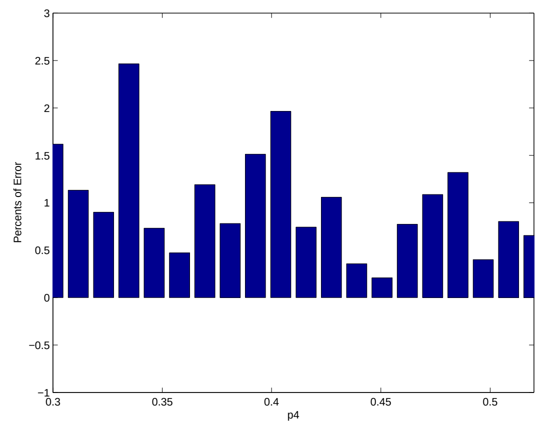

Fig. 5. Percent of Error vs. p_4

of the system is unchanged until p_4 reaches around 0.58. However the production rate begins to decrease when p_4 is bigger than 0.58. This is because p_4 less than 0.58, the bottleneck machine is M_2 and any parameter changes of the non-bottleneck machine does not influence the system production rate. However, if the p_4 is bigger than 0.58 the bottleneck machine becomes M_4 and the production rate decreases as the bottleneck machine deceases its capacity. Again, the analytical results are also the within the range of the confidence intervals evaluated from the simulation runs.

Figure 7 shows the percent of error of the production rate of the case 2. As shown in the figure all the errors remain within 3%. Notice that the analytical result tends to over estimate the production rate when M_2 is bottleneck, while it under estimate when M_4 is bottleneck. This behavior should be investigated in the future.

VI. CONCLUSION

In this paper, we introduced the analytical modeling and analysis of the re-entrant production line with two processing stages. We applied the existing decomposition equations for the multiple-part type production line and modified the decomposition equations to construct the re-entrant system. For verifications, the results from the analytical model is compared with the results from simulations runs. From the verification, we found that the analytical results were well matched with our intuitions and results from the simulation runs. In this paper the verifications were limited to the small re-entrant system with four machines and four buffers. For the next research step, we will extend the system to longer line with multiple re-entrant stages.

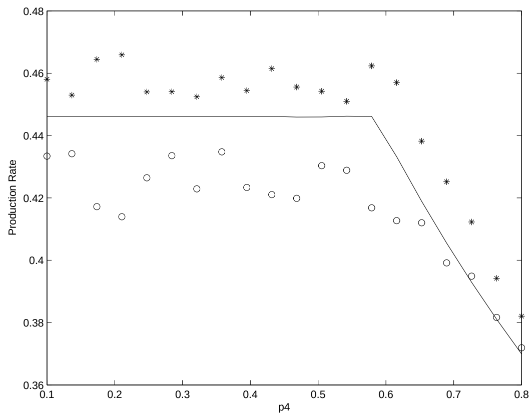

Fig. 6. Production rate vs. p_4 (r_4 fixed)

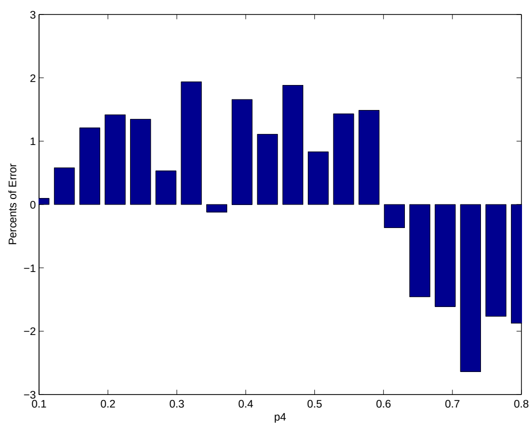

Fig. 7. Percent of Error vs. p_4

REFERENCES

[1] S. B. Gershwin. An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking. *Operations Research*, 35(2):291–305, 1987.

[2] S. B. Gershwin. Assembly/disassembly systems: An efficient decomposition algorithm for tree-structured networks. *IIE Transactions*, 23(4):302–314, 1991.

[3] Stanley B. Gershwin and Loren Werner. An approximate analytical method for evaluating the performance of closed loop flow systems with unreliable machines and finite buffers. *International Journal of Production Research*, 2006. To appear.

[4] Young Jae Jang and Stanley B. Gershwin. Modeling and analysis of two-part type manufacturing systems. In *MIT-Singapore Alliance Annual Symposium*, 2005.

[5] Nicola Maggio. An analytical method for evaluating the performance of closed loop production lines with unreliable machines and finite buffer. Master's thesis, Politecnico di Milano, 2000.