

Bottlenecks in Markovian Production Lines: A Systems Approach

S.-Y. Chiang, C.-T. Kuo, and S. M. Meerkov

Abstract—In this paper, a system-theoretic approach to bottlenecks in Markovian production lines is introduced and analyzed. The approach is based on the sensitivity of the system production rate to machines' reliability parameters. Using this approach, definitions of bottlenecks are introduced, methods for their identification are developed, and their implications for production automation and preventative maintenance are discussed.

Index Terms—Markovian statistics of breakdowns, production bottlenecks, production systems, unreliable machines.

I. INTRODUCTION

Production lines are sets of machines and material handling devices arranged in the consecutive order so as to produce a desired product. From a system-theoretic perspective, production lines are discrete-event systems. Investigation of fundamental laws that govern their behavior has implication for both production automation and preventative maintenance.

One of the problems that impedes the performance of production lines is the lack of machines' reliability: An unscheduled down-time of a machine may negatively affect the performance of all other machines, both up-stream and down-stream, blocking the former and starving the latter. Two basic models of machines reliability have been discussed in the literature: Bernoulli [1], [2] and Markovian [3]–[5]. Bernoulli model assumes that the status of a machine in each cycle (i.e., the time necessary to process one part) is determined by the process of Bernoulli trials. In Markovian model the state of a machine in a cycle is determined by a conditional probability, with the condition being the state of the machine in the previous cycle. This gives rise to a Markov process, which is why the term "Markovian production lines" is used.

Both models, Bernoulli and Markovian, reflect practical situations, however different ones. Bernoulli reliability model is more appropriate when down-time is small and comparable with the cycle time. This is often the case in assembly operations, where the down time is due to quality problems (see [1]). Markovian model reflects operations where the down-time is due to mechanical failures which could be much longer than the cycle time, [3]–[5]. In this paper we address the Markovian model.

Intuitively, *bottleneck (BN)* of a production line is understood as a machine that impedes the system performance in the strongest manner. Identification of BNs and their improvement, either by automation or work restructuring, is considered as one of the most important problems in current manufacturing environment. In fact, the

Manuscript received April 22, 1997; revised November 26, 1997. This work was supported by the U.S. National Science Foundation under Grant 9531694 and the Taiwan National Science Council under Grant NSC 87-2212-E-036-002. This paper was recommended for publication by Associate Editor N. Viswanadham and Editor P. B. Luh upon evaluation of the reviewers' comments.

S.-Y. Chiang and S. M. Meerkov are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA.

C.-T. Kuo is with the Department of Electrical Engineering, Tatung Institute of Technology, Taipei 104, Taiwan, R.O.C.

Publisher Item Identifier S 1042-296X(98)02097-7.

so-call Toyota Production System considers the continuous improvement as a part of three most important manufacturing programs: Just-in-Time (low inventories), Jidoka (quality), and Kaizen (continuous improvement) [6].

Unfortunately, not much is known, at least from the theoretical perspective, about BNs in production systems. Even the definition of a BN is unclear. For instance, some authors define the BN as the machine with the smallest isolation production rate (i.e. the production rate of the machine when no starvation and blockages are present). Others call the BN the machine with the largest inventory accumulated in front of it (see [7] and references therein). Both may not identify the machine that affects the bottom line, i.e. the system production rate, in the strongest possible manner (see [8]). This happens because the above definitions are local in nature and do not take into account the total *system* properties, such as the order of the machines in the production line, capacity of the material handling devices (buffers), etc. Therefore, a system-theoretic definition of a BN is necessary.

In the framework of Bernoulli lines, such a definition was introduced in [8], where a machine was referred to as the BN if the partial derivative of the system production rate with respect to machine's isolation production rate was the largest. Since the system production rate depends on all—the order of the machines in the line, their performance and reliability parameters, the buffer structure, etc.—this definition captures the system nature of the BN.

This definition, however, is not applicable to Markovian production lines. The reason is that, unlike the Bernoulli case, where the machine's production rate is defined by a single parameter, in the Markovian model the production rate is defined by two independent variables which characterize the up- and down-time. This makes it impossible to directly extend the Bernoulli case to the Markovian model. Therefore, the goal of this paper is to introduce definitions of BNs applicable to Markovian lines, suggest methods for their identification for the case of one- and two-machine production systems, and discuss implications of the results obtained for production automation and preventative maintenance.

The outline of the paper is as follows: In Section II, Markovian model of a production system is introduced, definitions of BNs are formulated, and the problems addressed are stated. Section III is devoted to the case of a single machine production system. Sections IV and V treat the two machines case. Finally, the conclusions are formulated in Section VI. The proofs are given in Appendices A and B.

II. DEFINITION AND PROBLEM FORMULATION

A. System Model

The following model of a production line is considered throughout this work.

- 1) The system consists of M machines arranged serially and $M - 1$ buffers separating each consecutive pair of machines.
- 2) Each machine m_i has two states: up and down. When up, the machine is capable of producing with the rate 1 part per unit of time (cycle); when the machine is down, no production takes place.
- 3) The up-time and the down-time of each machine m_i are random variables distributed exponentially with parameters p_i and r_i , respectively.

- 4) Each buffer b_i is characterized by its capacity, $N_i < \infty$, $1 \leq i \leq M - 1$.
- 5) Machine m_i is starved at time t if buffer b_{i-1} is empty at time t ; machine m_1 is never starved.
- 6) Machine m_i is blocked at time t if buffer b_i is full at time t ; machine m_M is never blocked.

A production line defined by 1)–6) will be denoted as $\{p_1, r_1, \dots, p_M, r_M, N_1, \dots, N_{M-1}\}$.

Remark 2.1: Due to assumption 3), the average up- and down time of the machines are

$$T_{up_i} = \frac{1}{p_i}; \quad T_{down_i} = \frac{1}{r_i}, \quad i = 1, \dots, M.$$

Therefore, the isolation production rate of each machine (i.e., the average number of parts produced per unit time if no starvation or blockage takes place) is

$$PR_{iso_i} = \frac{T_{up_i}}{T_{up_i} + T_{down_i}} = \frac{1}{1 + \frac{T_{down_i}}{T_{up_i}}}, \quad i = 1, \dots, M. \quad (2.1)$$

The isolation production rate is often referred to as machine efficiency and denoted as $e_i, i = 1, \dots, M$. Note that the machine efficiency depends only on the ratio of T_{up_i} and T_{down_i} , rather than on their absolute values. \square

Production rate of the line $\{p_1, r_1, \dots, p_M, r_M, N_1, \dots, N_{M-1}\}$ is the average number of parts produced by the last machine, m_M , per cycle. Given model 1)–6), the production rate, PR, is a function of all the machines and buffers parameters

$$PR = PR(p_1, r_1, p_2, r_2, \dots, p_M, r_M, N_1, \dots, N_{M-1}).$$

We use this function below to define the bottlenecks.

B. Bottleneck Definitions

Definition 2.1: Machine m_i is the *up-time bottleneck (UT-BN)* if

$$\frac{\partial PR}{\partial T_{up_i}} > \frac{\partial PR}{\partial T_{up_j}}, \quad j \neq i.$$

It is the *down-time bottleneck (DT-BN)* if

$$\left| \frac{\partial PR}{\partial T_{down_i}} \right| > \left| \frac{\partial PR}{\partial T_{down_j}} \right|, \quad j \neq i. \quad \square$$

Definition 2.2: Machine m_i is the *bottleneck (BN)* if it is both UT-BN and DT-BN. \square

Definition 2.3: Let m_i be the bottleneck machine. Then it is referred to as the *up-time preventative maintenance bottleneck (UTPM-BN)* if

$$\frac{\partial PR}{\partial T_{up_i}} > \left| \frac{\partial PR}{\partial T_{down_i}} \right|.$$

If the inequality is reversed, the bottleneck is referred to as the *down-time preventative maintenance bottleneck (DTPM-BN)*. \square

Remark 2.2: Definitions 2.1–2.3 are formulated in terms of the up- and down-time, rather than p_i and r_i of model 2)–6), because these are the variables used in practical situations on the factory floor. The differentiability of PR with respect to T_{up_i} and T_{down_i} is proved in Lemma A.1 of Appendix A. The absolute values of $\partial PR / \partial T_{down_i}$ is used because otherwise this number is negative: increase in T_{down} leads to a decrease of PR. \square

The meaning of UT-BN and DT-BN is clear from Definition 2.1: A machine is UT-BN (or DT-BN) if an increase of its up-time (respectively, a decrease of its down-time) leads to the largest increase

of the system production rate. The meaning of BN of Definition 2.2 is also obvious: A machine is the BN if both its up-time and its down-time are the most critical for the system performance. Since, in some instances, the down-time of a machine is due to lapses in the performance of manual operators, rather than machines breakdown, the identification of DT-BN and/or BN provides guidance for development of production automation.

The meaning of the *up-time preventative maintenance bottleneck* and the *down-time preventative maintenance bottleneck* of Definition 2.3 needs an explanation: Preventative maintenance, as part of the total production maintenance, leads to both an increase of the up-time and a decrease of the down-time of automated machines. Some of the preventative maintenance measures affect more the up-time and the others the down-time. We refer to them as up-time preventative maintenance (UTPM) and down-time preventative maintenance (DTPM), respectively. Which one of these measures should be emphasized? If it is determined that the BN is, in fact, UTPM-BN, then the former should have the priority in skilled trades job assignment; otherwise the latter should have the precedence. Thus, the classification of the BN in either UTPM or DTPM has an impact on planning actions that lead to the most efficient system improvement.

The following is a direct consequence of Definitions 2.1–2.3: Production line 1)–6) always has an UT-BN and DT-BN; it may or may not have a BN; generically, the BN is of a unique nature, either UTPM-BN or DTPM-BN. Even if a line does not have a BN, the identification of UT-BN and DT-BN may be used for the development of automation and skilled trades assignment in a manner outlined above.

C. Problem Formulation

The goal of this work is to derive a tool for identification of the bottlenecks defined above. Unfortunately, direct identification of BNs using Definitions 2.1–2.3 is, in many cases, impossible. The reason is two fold: First, the derivatives of the production rate involved cannot be measured on the factory floor during the normal system operation. Second, in most cases they cannot be calculated analytically as well since even the calculation of the production rate for system with more than two machines is impossible (see [3]–[5]), let alone the calculation of its derivatives. Therefore, the tools sought have to be indirect ones. More specifically, we are seeking BN identification tools that are based on either the data available on the factory floor through real time measurements (such as average up- and down-time, starvation and blockage time, etc.) or on the data that can be constructively calculated using the machines and buffers parameters (r_i, p_i , and N_i). We refer to these tools as *Bottleneck Indicators*. The problems, then, addressed in this paper are:

Problem 1: Given a production systems defined by 1)–6), derive Bottleneck Indicators for UT- and DT-BN identification. This, of course, will lead to BN identification as well.

Problem 2: Given a production systems defined by 1)–6), derive Bottleneck Indicators for UTPM- and DTPM-BN identification.

Solution of these problems for the simplest cases of one- and two-machine systems is given below. An extension to the case of $M > 2$ case is a subject of the future work.

III. ONE MACHINE CASE

If the system consists of a single machine, Definition 2.3 is the only applicable one, i.e., either UTPM-BN or DTPM-BN could exist. When each of these takes place is quantified as follows:

Theorem 3.1: A single machine defined by assumptions 2)–3) is UTPM-BN if $T_{up} < T_{down}$; it is DTPM-BN if $T_{down} < T_{up}$.

Proof: Follows immediately from (2.1) since

$$\left| \frac{\partial \text{PR}}{\partial T_{\text{down}}} \right| = \frac{T_{\text{up}}}{(T_{\text{up}} + T_{\text{down}})^2}$$

and

$$\frac{\partial \text{PR}}{\partial T_{\text{up}}} = \frac{T_{\text{down}}}{(T_{\text{up}} + T_{\text{down}})^2}.$$

■

Based on the above, we arrive at the following

Bottleneck Indicator 3.1: The smaller between the average up-time and average down-time of a machine defines its nature as the BN: If $T_{\text{down}} < T_{\text{up}}$, the primary attention of the preventative maintenance and automation should be given to the further decrease of the down-time; if $T_{\text{up}} < T_{\text{down}}$, the attention should be concentrated on the increase of the up-time. □

Since in most practical situations $T_{\text{down}} < T_{\text{up}}$, the above Indicator basically states that reduction of the down-time is more efficient than a comparable increase of the up-time.

IV. TWO MACHINES CASE: EQUAL ISOLATION PRODUCTION RATES

Consider system 1)–6) with two machines, i.e., system $\{T_{\text{up}1}, T_{\text{down}1}, T_{\text{up}2}, T_{\text{down}2}, N_1\}$. Assume that

$$\frac{T_{\text{up}1}}{T_{\text{down}1}} = \frac{T_{\text{up}2}}{T_{\text{down}2}} \quad (4.1)$$

i.e., the efficiency of both machines are identical, however, the up- and down-time for each machine may be different. A question arises: Which one is the BN? The answer is as follows.

Theorem 4.1: Consider the production line $\{T_{\text{up}1}, T_{\text{down}1}, T_{\text{up}2}, T_{\text{down}2}, N_1\}$ and assume that (4.1) holds. Then m_1 is the BN if $T_{\text{down}1} < T_{\text{down}2}$; if the inequality is reversed, m_2 is the BN.

Proof: See Appendix A.

It is well known that, given a constant ratio between $T_{\text{up}i}$ and $T_{\text{down}i}$, the machine with the longer up- and down-time is more detrimental to the system's production rate than that with a shorter up- and down-time (see, for instance, [5]). In view of this property, one might think that the BN is the machine with the longer up- and down-time. This, however, is not true, as the above Theorem states. The reason is that an improvement of the machine with a shorter up- and down-time leads to a better utilization of the disturbance attenuation capabilities of the buffer than a comparable improvement of the machine with a longer up- and down-time. Therefore, an improvement of the "better" machine is the best for the system as a whole.

Assume now that (4.1) holds and, in addition

$$\frac{T_{\text{up}i}}{T_{\text{up}j}} = \frac{T_{\text{down}i}}{T_{\text{down}j}} = k > 1, \quad i, j = 1, 2, \quad i \neq j \quad (4.2)$$

i.e., m_j is the BN. Is it UTPM-BN or DTPM-BN? The answer is in the following.

Theorem 4.2: Under assumption (4.2), m_j is DTPM-BN if $T_{\text{down}j} < T_{\text{up}j}$. If $T_{\text{down}j}$ is sufficiently larger than $T_{\text{up}j}$, so that

$$\frac{T_{\text{down}j}}{T_{\text{up}j}} > \frac{k+1}{k}$$

m_j is UTPM-BN.

Proof: See Appendix A. ■

The above result is qualitatively in agreement with Theorem 3.1: The smaller between the up- and down-time of the BN defines its nature as far as the preventative maintenance is concerned.

From Theorems 4.1 and 4.2 follows

Bottleneck Indicator 4.1: In a production line with two machines of equal efficiency, the machine with the smaller down-time is the BN. If the down-time of this machine is smaller than its up-time, preventative maintenance and automation should be directed toward the decrease of the down-time. If the down-time is sufficiently larger than the up-time, preventative maintenance and automation should be directed toward the increase of the up-time. □

Assume finally that (4.1) holds and, in addition

$$T_{\text{up}1} = T_{\text{up}2} = T_{\text{up}}, \quad T_{\text{down}1} = T_{\text{down}2} = T_{\text{down}} \quad (4.3)$$

i.e., both machines are identical as far as their reliability is concerned. In this situation, obviously, both machines are BNs but of which kind? The answer is given below:

Theorem 4.3: Consider $\{T_{\text{up}1}, T_{\text{down}1}, T_{\text{up}2}, T_{\text{down}2}, N_1\}$ and assume that (4.3) holds. Then both machines are DTPM-BN if

$$\frac{T_{\text{up}}}{T_{\text{up}} + T_{\text{down}}} > 0.5$$

both machines are UTPM-BN if

$$\frac{T_{\text{up}}}{T_{\text{up}} + T_{\text{down}}} < 0.4315.$$

Proof: See Appendix A. ■

As it has been shown in Section III for a single machine case, the threshold for the switch from DTPM-BN to UTPM-BN is 0.5. In the two identical machines case, the threshold is a function of the buffer capacity, N_1 , i.e., has the form $\tau(N_1)$. Theorem 4.3 establishes, therefore, that $0.4315 < \tau(N_1) < 0.5$. When N_1 tends to infinity, $\tau(N_1)$ tends to 0.5. The minimum threshold occurs for $N_1 = 0.8T_{\text{down}}$ (see the proof of Theorem 4.3).

From Theorem 4.3 follows

Bottleneck Indicator 4.2: In a production line consisting of two machines with identical reliability characteristics, the primary attention of the preventative maintenance and automation should be given to the decrease of the down-time if $T_{\text{down}} < T_{\text{up}}$; if $T_{\text{up}} < 0.759 T_{\text{down}}$, attention should be concentrated on the increase of the up-time.

In most practical situations, the isolation production rate of the machines is greater than 0.5. Therefore, taking into account Theorems 3.1, 4.2, and 4.3, we concentrate below on the identification of DT-BNs.

V. TWO MACHINES CASE: UNEQUAL ISOLATION PRODUCTION RATES

Consider again $\{T_{\text{up}1}, T_{\text{down}1}, T_{\text{up}2}, T_{\text{down}2}, N_1\}$ and assume that

$$\frac{T_{\text{up}1}}{T_{\text{down}1}} \neq \frac{T_{\text{up}2}}{T_{\text{down}2}} \quad (5.1)$$

i.e., the machines are of unequal efficiency. Is the machine with the smallest efficiency the BN? The answer is: not necessarily. An example in Fig. 1 illustrates this point. In this example, the numbers in the circles represent the machines' efficiency and the number in the rectangle is the buffer capacity. The first two rows of the numbers below the machines are up- and down-time, respectively. The third row is a numerical estimate of the partial derivative of the production rate with respect to the down-time. As it follows from this figure, although m_1 is more efficient than m_2 , it is, nevertheless, the DT-BN of the system.

To identify the DT-BN in the case of machines with unequal efficiency, introduce the following concepts:

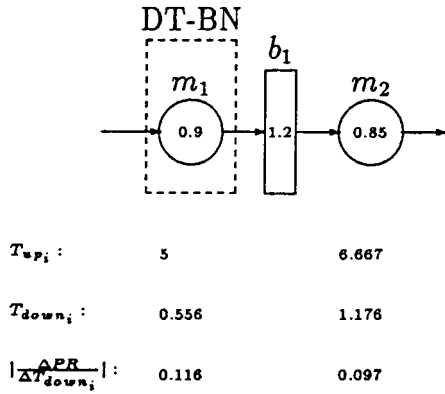


Fig. 1. The bottleneck example.

Definition 5.1: Machine m_i is said to be *blocked in the manufacturing sense* at time t if it is up at time t , buffer b_i is full at time t , and machine m_{i+1} fails to take parts at time t . The probability of manufacturing blockage, mb_i , is defined as

$$mb_i = \text{Prob}(\{m_i \text{ is up at time } t\} \cap \{b_i \text{ is full at time } t\} \cap \{m_{i+1} \text{ fails to take parts from } b_i \text{ at time } t\}).$$

Definition 5.2: Machine m_i is said to be *starved in the manufacturing sense* at time t if it is up at time t , machine m_{i-1} fails to put parts into buffer b_{i-1} at time t , and buffer b_{i-1} is empty at time t . The probability of manufacturing starvation, ms_i , is defined as

$$ms_i = \text{Prob}(\{m_{i-1} \text{ fails to put parts into } b_{i-1} \text{ at time } t\} \cap \{b_{i-1} \text{ is empty at time } t\} \cap \{m_i \text{ is up at time } t\}).$$

□

In the case of Bernoulli lines, ms_i and mb_i were sufficient to characterize the BNs (see [8]). It turns out, however, that in Markovian case they should be modified appropriately to serve this purpose. Nevertheless, they play an important role in BN identification and, therefore, are characterized, in terms of the parameters of model 1)-6), below.

Lemma 5.1: For the serial production line 1)–6) with $M = 2$

$$\begin{aligned} mb_1 &= e_1 Q(p_2, r_2, p_1, r_1, N_1) \\ ms_2 &= e_2 Q(p_1, r_1, p_2, r_2, N_1) \end{aligned} \quad (5.2)$$

where $p_i = 1/T_{up_i}$ and $r_i = 1/T_{down_i}$, $i = 1, 2$, Q is seen in (5.3), shown at the bottom of the page, and

$$\begin{aligned} \phi &= \frac{e_1(1 - e_2)}{e_2(1 - e_1)} \\ \beta &= \frac{(r_1 + r_2 + p_1 + p_2)(p_1 r_2 - p_2 r_1)}{(r_1 + r_2)(p_1 + p_2)}. \end{aligned}$$

Proof: See Appendix B. ■

Lemma 5.2: For the serial production line 1)–6) with $M = 2$ under assumption (4.1)

$$mb_1 = ms_2. \quad (5.4)$$

Proof: See Appendix B. ■

The values of mb_1 and ms_2 , calculated according Lemma 5.1, play a crucial role in DT-BN identification in production lines with unequal machines efficiency. Specifically,

Bottleneck Indicator 5.1: If $mb_1 T_{up_1} T_{down_1} < ms_2 T_{up_2} T_{down_2}$, m_1 is the DT-BN. If $mb_1 T_{up_1} T_{down_1} > ms_2 T_{up_2} T_{down_2}$, DT-BN is m_2 .

Remark 5.1: Taking into account Lemma 5.2, it is easy to see that Bottleneck Indicator 5.1 is a generalization of Bottleneck Indicator 4.1 and the Bottleneck Indicator of [8]. Indeed, in the case of equal machines' efficiency, machine m_1 is the BN if $T_{up_1} T_{down_1} < T_{up_2} T_{down_2}$; if the inequality is reversed, m_2 is the BN. In the case of Bernoulli lines [8], m_1 is the BN if $mb_1 < ms_2$; if the inequality is reversed, m_2 is the BN. Thus, Bottleneck Indicator 5.1 is a generalization of the above two results. ■

Numerical Justification: At this time, Bottleneck Indicator 5.1 has been justified only numerically. Two typical examples are given in Fig. 2. In this figure, the first two rows of numbers below the machines are the average up-time and down-time, respectively. The other three rows of numbers show the values of $mb_i T_{up_i} T_{down_i}$, $ms_i T_{up_i} T_{down_i}$, and $|\Delta PR / \Delta T_{down_i}|$. The finite differences $\Delta PR / \Delta T_{down_i}$, $\forall i$, are numerical estimates of $\partial PR / \partial T_{down_i}$ with the step $\Delta T_{down_i} = 0.05 \cdot T_{down_i}$. Then, from Bottleneck Indicator 5.1, the bottlenecks in Fig. 2(a) and (b) are machines m_2 and m_1 , respectively. In most systems investigated the bottlenecks identified using Bottleneck Indicator 5.1 and $|\Delta PR / \Delta T_{down_i}|$ were the same. However, several counterexamples were also found. To illustrate the region where Bottleneck Indicator 5.1 does not work (Fig. 3), we use the parameters of the system in Fig. 2(a). In Fig. 3(a), T_{down_2} is changed from 1.176–0.08 and the bottleneck machine is switched from m_2 to m_1 . As it follows from this figure, the range where Bottleneck Indicator 5.1 does not work is quite small. In Fig. 3(b), the efficiency of m_2 is used as the horizontal axis. We also obtain a similar result (i.e., the efficiency region where the Bottleneck Indicator 5.1 does not work is very small [about 0.01]). Based on the above, we conclude that it can be used as a tool for the DT-BN identification.

VI. CONCLUSION

Bottlenecks in Markovian production lines can be defined as partial derivatives of the system production rate with respect to machines' up- and down-time. Three types of BNs are introduced:

- 1) up- and down-time BNs;
- 2) BNs;
- 3) UTPM- and DTPM-BNs.

Identification of each has implications for both the nature of preventative maintenance and production automation.

In "balanced" two-machine lines (i.e., in lines where machines have identical efficiency) the machine with the smaller down-time is the BN. In "unbalanced" two-machine lines, the DT-BN is the machine

$$Q(p_1, r_1, p_2, r_2, N_1) = \begin{cases} \frac{(1 - e_1)(1 - \phi)}{1 - \phi e^{-\beta N_1}}, & \frac{p_1}{r_1} \neq \frac{p_2}{r_2} \\ \frac{p_1(p_1 + p_2)(r_1 + r_2)}{(p_1 + r_1)[(p_1 + p_2)(r_1 + r_2) + p_2 r_1(p_1 + p_2 + r_1 + r_2)N_1]}, & \frac{p_1}{r_1} = \frac{p_2}{r_2} \end{cases} \quad (5.3)$$

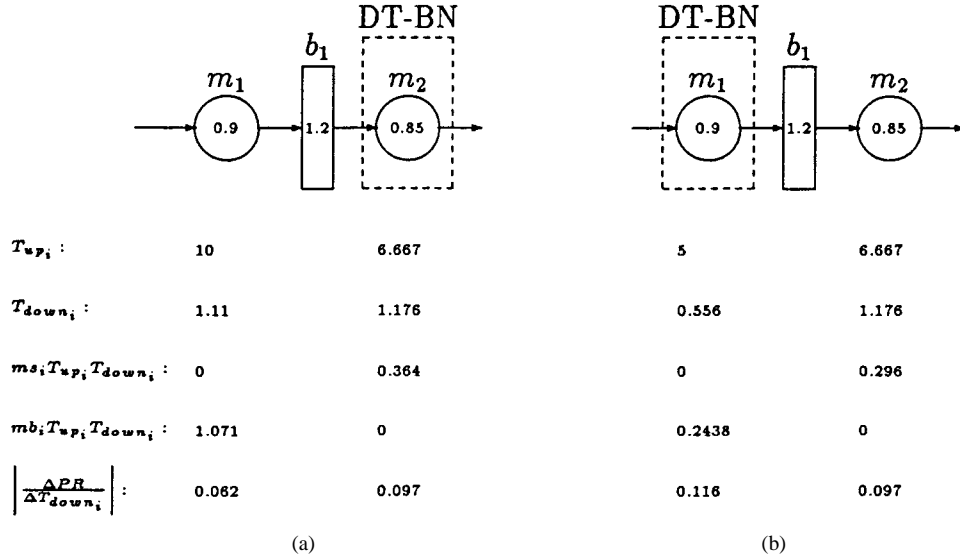
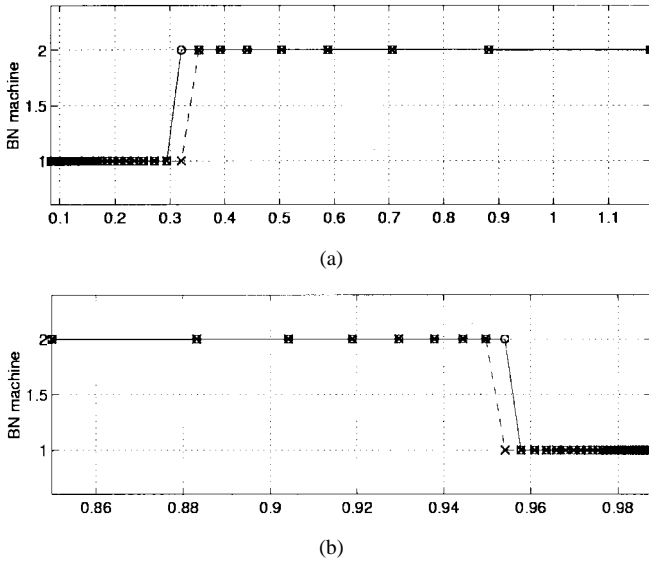


Fig. 2. Illustration of Bottleneck Indicator 5.1.

Fig. 3. Comparison of Bottleneck Indicator 5.1 with $\max(\partial PR/\partial T_{down})$: $\times\times\times$ —BN identified by BN Indicator 5.1, $\circ\circ\circ\circ$ —BN identified by $\max(\partial PR/\partial T_{down})$.

with the smallest value of $\alpha T_{up} T_{down}$, where α is the probability of blockage for the first machine and the probability of starvation for the

second. Current work is centered on extensions of the results obtained to systems with arbitrary number of machines and on applications in the automotive industry.

APPENDIX A PROOFS FOR SECTION IV

Due to space limitations, we present here and in Appendix B only the main steps of the proofs. The details can be found in [10].

The proof of Theorem 4.1 requires the expression for the production rate of two-machine system and the property of its differentiability with respect to the machines reliability parameters. The former follows from [9] in the form (A.1), as shown at the bottom of the page, where

$$\beta = \frac{(r_1 + r_2 + p_1 + p_2)(p_1 r_2 - p_2 r_1)}{(r_1 + r_2)(p_1 + p_2)}. \quad (\text{A.2})$$

The latter is established by the following.

Lemma A.1: For all $0 \leq N_1 < \infty$, production rate given by (A.1) is differentiable with respect to the remaining arguments p_i, r_i, T_{up_i} , and $T_{down_i}, i = 1, 2$.

Proof— Step 1: For the case of $p_1/r_1 = p_2/r_2$, using straightforward, but tedious, calculations, we show that the right derivative of PR with respect to p_1 is expressed as (A.3) as shown at the bottom of the page, where

$$\text{PR} = \begin{cases} \frac{r_1 r_2}{(p_1 + r_1)(p_2 + r_2)} \left[\frac{p_1(p_2 + r_2) - p_2(p_1 + r_1)e^{-\beta N_1}}{p_1 r_2 - p_2 r_1 e^{-\beta N_1}} \right], & \frac{p_1}{r_1} \neq \frac{p_2}{r_2} \\ \frac{r_2^2(r_1 + r_2) + N_1 r_1 r_2 (p_2 + r_2)^2}{(p_2 + r_2)^2 [r_1 + r_2 + N_1 r_1 (p_2 + r_2)]}, & \frac{p_1}{r_1} = \frac{p_2}{r_2} \end{cases} \quad (\text{A.1})$$

$$\begin{aligned} & \lim_{\Delta p \rightarrow 0^+} \frac{\text{PR}(p_1 + \Delta p, r_1, p_2, r_2, N_1) - \text{PR}(p_1, r_1, p_2, r_2, N_1)}{\Delta p} \\ &= \frac{r_1 r_2 \{ [n_1 g_1 - p_2(p_1 + r_1) a_1 g_1 - f_1(d_1 - p_2 r_1 a_1)](p_1 + r_1) - f_1 g_1 \}}{(p_1 + r_1)^2 (p_2 + r_2) g_1^2} \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned}
f_1 &= 1 + p_2(p_1 + r_1) \frac{r_1 + r_2 + p_1 + p_2}{(r_1 + r_2)(p_1 + p_2)} N_1 \\
g_1 &= 1 + p_2 r_1 \frac{r_1 + r_2 + p_1 + p_2}{(r_1 + r_2)(p_1 + p_2)} N_1 \\
n_1 &= N_1 p_2 \\
&\cdot \frac{(p_1 + p_2)(r_1 + r_2 + p_1 + p_2) - (p_1 + r_1)(r_1 + r_2)}{(r_1 + r_2)(p_1 + p_2)^2} \\
d_1 &= \frac{-p_2 r_1 N_1}{(p_1 + p_2)^2} \\
a_1 &= \frac{(r_1 + r_2 + p_1 + p_2)^2 r_2 N_1^2}{2(r_1 + r_2)^2 (p_1 + p_2)^2}. \tag{A.4}
\end{aligned}$$

Step 2: Analogously, the left derivative PR is (A.5), as shown at the bottom of the page, where f_1, g_1, n_1, d_1 , and a_1 are given by (A.4).

Step 3: From (A.3) and (A.5), it follows that the production rate PR is differentiable with respect to p_1 . Analogously, we show that PR is differentiable with respect to with respect to r_1, p_2 , and r_2 .

Step 4: Since

$$\begin{aligned}
\frac{\partial \text{PR}}{\partial T_{\text{up}_i}} &= \frac{\partial \text{PR}}{\partial p_i} \frac{\partial p_i}{\partial T_{\text{up}_i}} = \frac{\partial \text{PR}}{\partial p_i} (-p_i^2), \quad i = 1, 2 \\
\frac{\partial \text{PR}}{\partial T_{\text{down}_i}} &= \frac{\partial \text{PR}}{\partial r_i} \frac{\partial r_i}{\partial T_{\text{down}_i}} = \frac{\partial \text{PR}}{\partial r_i} (-r_i^2), \quad i = 1, 2
\end{aligned} \tag{A.6}$$

it follows that the production rate PR is differentiable with respect to T_{up_i} and T_{down_i} , $i = 1, 2$.

Step 5: For $p_1/r_1 \neq p_2/r_2$, Steps 1–4 are repeated to establish the differentiability of PR in this case as well.

Proof of Theorem 4.1— Step 1: From (A.1), we show (A.7), as seen on the bottom of the page, where

$$\begin{aligned}
f_i &= 1 + p_j(p_i + r_i) \frac{r_1 + r_2 + p_1 + p_2}{(r_1 + r_2)(p_1 + p_2)} N_1 \\
g_i &= 1 + p_j r_i \frac{r_1 + r_2 + p_1 + p_2}{(r_1 + r_2)(p_1 + p_2)} N_1 \\
n_i &= N_1 p_j \\
&\cdot \frac{(p_1 + p_2)(r_1 + r_2 + p_1 + p_2) - (p_i + r_i)(r_1 + r_2)}{(r_1 + r_2)(p_1 + p_2)^2} \\
d_i &= \frac{-p_j r_i N_1}{(p_1 + p_2)^2} \\
a_i &= \frac{(r_1 + r_2 + p_1 + p_2)^2 r_j N_1^2}{2(r_1 + r_2)^2 (p_1 + p_2)^2} \\
i, j &= 1, 2, \quad i \neq j. \tag{A.8}
\end{aligned}$$

Step 2: Under assumption (4.1), from the above expressions and (A.6), we show that

$$\begin{aligned}
\frac{\partial \text{PR}}{\partial T_{\text{up}_1}} - \frac{\partial \text{PR}}{\partial T_{\text{up}_2}} &= \frac{p_2^2 r_1 r_2 k (k - 1)}{2(p_1 + r_1)^2 (p_2 + r_2) (r_1 + r_2)^2 (p_1 + p_2)^2 g_1^2} \\
&\cdot \{N_1^2 p_1 p_2 r_1 (p_2 + r_2) (r_1 + r_2 + p_1 + p_2)^2 \\
&+ 2(r_1 + r_2)^2 (p_1 + p_2)^2 + 2N_1 p_2 (p_2 + r_2) \\
&\cdot [p_1 (r_1 + r_2) (p_2 + r_2) (k + 1) + r_1 (p_1 + p_2) \\
&\cdot (r_1 + r_2 + p_1 + p_2)]\} \tag{A.9}
\end{aligned}$$

where $k = T_{\text{up}_2}/T_{\text{up}_1} = T_{\text{down}_2}/T_{\text{down}_1} = p_1/p_2 = r_1/r_2$.

If $k > 1$, (i.e., the up- and down-time of m_1 are shorter than those of m_2), then, since $g_1 > 0$,

$$\frac{\partial \text{PR}}{\partial T_{\text{up}_1}} > \frac{\partial \text{PR}}{\partial T_{\text{up}_2}}. \tag{A.10}$$

Thus, the machine with the shorter up- and down-time is the UP-BN.

Step 3: Steps 1 and 2 are repeated for T_{down_i} , $i = 1, 2$, to obtain

$$\begin{aligned}
\left| \frac{\partial \text{PR}}{\partial T_{\text{down}_1}} \right| - \left| \frac{\partial \text{PR}}{\partial T_{\text{down}_2}} \right| &= \frac{p_2 r_1 r_2^2 k (k - 1)}{2(p_1 + r_1)^2 (p_2 + r_2) (r_1 + r_2)^2 (p_1 + p_2)^2 g_1^2} \\
&\cdot \{N_1^2 p_1 p_2 r_1 (p_2 + r_2) (r_1 + r_2 + p_1 + p_2)^2 \\
&+ 2(r_1 + r_2)^2 (p_1 + p_2)^2 + 2N_1 p_2 (p_1 + p_2) (r_1 + r_2) \\
&\cdot [r_1 (p_1 + p_2) + (r_1 + r_2) (2r_1 + p_1)]\}. \tag{A.11}
\end{aligned}$$

If $k > 1$, we have

$$\left| \frac{\partial \text{PR}}{\partial T_{\text{down}_1}} \right| > \left| \frac{\partial \text{PR}}{\partial T_{\text{down}_2}} \right| \tag{A.12}$$

i.e., the machine with the shorter up- and down-time is the DT-BN.

Step 4: From (A.10) and (A.12), it follows that the BN is the machine with the shorter up- and down-time.

Proof of Theorem 4.2— Step 1: From Theorem 4.1 and Assumption (4.2), we have $\partial \text{PR}/\partial T_{\text{up}_j} > \partial \text{PR}/\partial T_{\text{up}_i}$, and $|\partial \text{PR}/\partial T_{\text{down}_j}| > |\partial \text{PR}/\partial T_{\text{down}_i}|$, $i, j = 1, 2, i \neq j$, i.e., machine j is the BN.

$$\begin{aligned}
&\lim_{\Delta p \rightarrow 0} \frac{\text{PR}(p_1 + \Delta p, r_1, p_2, r_2, N_1) - \text{PR}(p_1, r_1, p_2, r_2, N_1)}{\Delta p} \\
&= \frac{r_1 r_2 \{ [n_1 g_1 - p_2 (p_1 + r_1) a_1 g_1 - f_1 (d_1 - p_2 r_1 a_1)] (p_1 + r_1) - f_1 g_1 \}}{(p_1 + r_1)^2 (p_2 + r_2) g_1^2} \tag{A.5}
\end{aligned}$$

$$\frac{\partial \text{PR}}{\partial p_i} = \frac{r_i r_j \{ [n_i g_i - p_j (p_i + r_i) a_i g_i - f_i (d_i - p_j r_i a_i)] (p_i + r_i) - f_i g_i \}}{(p_j + r_j) (p_i + r_i)^2 g_i^2} \tag{A.7}$$

$i = 1, 2, \quad i \neq j$

TABLE I
SUMMARY PROBABILITIES AND DENSITIES OF BUFFER OCCUPANCY FOR TWO MACHINES-ONE BUFFER SYSTEM (β IS GIVEN IN (5.3) AND $\kappa = [p_1 p_2 r_1 r_2 (r_1 + r_2 + p_1 + p_2)(p_1 r_2 - p_2 r_1)] / [(p_1 + p_2)^2 (p_1 + r_1)(p_2 + r_2)(p_1 r_2 - p_2 r_1 e^{-\beta N_1})]$)

$\alpha_1 \alpha_2$	$Y_{1;\alpha_1 \alpha_2}(0)$	$X_{1;\alpha_1 \alpha_2}(h_1)$	$Y_{1;\alpha_1 \alpha_2}(N_1)$
uu	$\kappa \frac{(p_1+p_2)}{p_2(r_1+r_2+p_1+p_2)}$	$\kappa e^{-\beta h_1}$	$\kappa \frac{(p_1+p_2)}{p_1(r_1+r_2+p_1+p_2)} e^{-\beta N_1}$
ud	0	$\kappa \frac{p_1+p_2}{r_1+r_2} e^{-\beta h_1}$	$\kappa \frac{(p_1+p_2)^2}{p_1 r_2 (r_1+r_2+p_1+p_2)} e^{-\beta N_1}$
du	$\kappa \frac{(p_1+p_2)^2}{p_2 r_1 (r_1+r_2+p_1+p_2)}$	$\kappa \frac{p_1+p_2}{r_1+r_2} e^{-\beta h_1}$	0
dd	$\kappa \frac{(p_1+p_2)^2}{r_1(r_1+r_2)(r_1+r_2+p_1+p_2)}$	$\kappa \left(\frac{p_1+p_2}{r_1+r_2}\right)^2 e^{-\beta h_1}$	$\kappa \frac{(p_1+p_2)^2}{r_2(r_1+r_2)(r_1+r_2+p_1+p_2)} e^{-\beta N_1}$

Step 2: To determine under what conditions machine j is DTPM-BN or UTPM-BN, we show that

$$\begin{aligned} & \left| \frac{\partial \text{PR}}{\partial T_{\text{down}j}} \right| - \frac{\partial \text{PR}}{\partial T_{\text{up}j}} \\ &= \frac{r_j r_i k^2}{(p_j + r_j)^2 (p_i + r_i) g_j^2 r_i (k+1)^2} \\ & \cdot \{p_i + r_i (k+1)^2 (r_i - p_i) + 2p_i r_i k (p_i + r_i) \\ & \cdot [k(r_i - p_i) + r_i] N_1 + 0.5 p_i k^2 (p_i + r_i)^3 \\ & \cdot (r_i - p_i) N_1^2\}, \quad i, j = 1, 2, \quad i \neq j \end{aligned} \quad (\text{A.13})$$

where g_j is given in (A.8) and $k = p_j/p_i = r_j/r_i$.

Step 3: From the above expressions, if $r_i > p_i$ (i.e., $r_j > p_j$ or $T_{\text{up}j} > T_{\text{down}j}$), then we have

$$\left| \frac{\partial \text{PR}}{\partial T_{\text{down}j}} \right| > \frac{\partial \text{PR}}{\partial T_{\text{up}j}} \quad (\text{A.14})$$

i.e., m_j is DTPM-BN. If $r_i < p_i$ and $k(r_i - p_i) + r_i < 0$ which implies that $T_{\text{down}i} > T_{\text{up}i}$ (i.e., $T_{\text{down}j} > T_{\text{up}j}$) and $(T_{\text{down}i}/T_{\text{up}i}) > (k+1)/k$ (i.e., $(T_{\text{down}j}/T_{\text{up}j}) > (k+1)/k$), we have

$$\left| \frac{\partial \text{PR}}{\partial T_{\text{down}j}} \right| < \frac{\partial \text{PR}}{\partial T_{\text{up}j}}. \quad (\text{A.15})$$

It follows that if $(T_{\text{down}j}/T_{\text{up}j}) > (k+1)/k$, m_j is UTPM-BN. ■

Proof of Theorem 4.3— Step 1: Under the Assumption (4.3), from (A.1), we show that

$$\text{PR} = \frac{2T_{\text{up}}[2T_{\text{up}}^2 T_{\text{down}} + N_1(T_{\text{up}} + T_{\text{down}})^2]}{2(T_{\text{up}} + T_{\text{down}})^2[2T_{\text{up}} T_{\text{down}} + N_1(T_{\text{up}} + T_{\text{down}})]}. \quad (\text{A.16})$$

Step 2: From the above expression, it follows:

$$\begin{aligned} & \frac{\partial \text{PR}}{\partial T_{\text{up}}} - \left| \frac{\partial \text{PR}}{\partial T_{\text{down}}} \right| \\ &= \frac{T_{\text{up}} + T_{\text{down}}}{(T_{\text{up}} + T_{\text{down}})^4 [2T_{\text{up}} T_{\text{down}} + N_1(T_{\text{up}} + T_{\text{down}})]^2} \\ & \cdot [N_1^2 (T_{\text{up}} + T_{\text{down}})^3 (T_{\text{down}} - T_{\text{up}}) + 4N_1 T_{\text{up}}^2 \\ & \cdot T_{\text{down}} (T_{\text{up}} + T_{\text{down}}) (T_{\text{down}} - 2T_{\text{up}}) + 8T_{\text{up}}^3 \\ & \cdot T_{\text{down}}^2 (T_{\text{down}} - T_{\text{up}})]. \end{aligned} \quad (\text{A.17})$$

Step 3: The threshold τ (i.e., the efficiency such that $|\partial \text{PR}/\partial T_{\text{down}}| = \partial \text{PR}/\partial T_{\text{up}}$) is defined by

$$\frac{\partial \text{PR}}{\partial T_{\text{up}}} - \left| \frac{\partial \text{PR}}{\partial T_{\text{down}}} \right| = 0. \quad (\text{A.18})$$

Step 4: From the above expression, for the minimum τ , we show that

$$N_1 = \frac{-2T_{\text{up}}^2 T_{\text{down}} (T_{\text{down}} - 2T_{\text{up}})}{(T_{\text{up}} + T_{\text{down}})^2 (T_{\text{down}} - T_{\text{up}})}. \quad (\text{A.19})$$

Step 5: From (A.18) and (A.19), it follows that the threshold values of T_{up} and T_{down} are defined by the following equation:

$$T_{\text{down}}^3 - 1.5T_{\text{up}} T_{\text{down}}^2 + T_{\text{up}}^2 T_{\text{down}} - T_{\text{up}}^3 = 0. \quad (\text{A.20})$$

Solving this equation by Matlab, we obtain

$$T_{\text{down}} = 1.31718264650677 T_{\text{up}}. \quad (\text{A.21})$$

Moreover, substituting (A.21) into (A.19), we have

$$N_1 = 0.8 T_{\text{down}}.$$

Step 6: Therefore, from (A.21), if $T_{\text{up}}/(T_{\text{up}} + T_{\text{down}}) < 0.4315$, both machines are UT-BNs. From (A.17), if $T_{\text{up}} > T_{\text{down}}$ (i.e., $T_{\text{up}}/(T_{\text{up}} + T_{\text{down}}) > 0.5$), both machines are DT-BNs. ■

APPENDIX B PROOFS FOR SECTION V

Consider the system defined by assumption 1)–6) with $M = 2$. Introduce the stationary probability distribution, $Y_{1;\alpha_1 \alpha_2}(h_1)$, as follows:

$$\begin{aligned} Y_{1;\alpha_1 \alpha_2}(h_1) = & \text{Prob}\{\text{buffer } b_1 \text{ contains } h_1 \text{ parts} \\ & \text{and machines } m_1 \text{ and } m_2 \text{ are in} \\ & \text{the states of } \alpha_1 \text{ and } \alpha_2\} \end{aligned}$$

where $0 \leq h_1 \leq N_1$, and $\alpha_i = u$ (i.e., m_i is up) or d (i.e., m_i is down), $i = 1, 2$. The PR of this two machines-one buffer system can be calculated using the following:

Lemma B.1: For the serial production line 1)–6) with $M = 2$

$$\begin{aligned}
 \text{PR} &= e_2 - Y_{1;du}(0) \\
 &= e_1 - Y_{1;ud}(N_1) \\
 &= e_2[1 - Q(p_1, r_1, p_2, r_2, N_1)] \\
 &= e_1[1 - Q(p_2, r_2, p_1, r_1, N_1)]
 \end{aligned} \tag{B.1}$$

where function $Q(x_1, y_1, x_2, y_2, N_1)$ is defined in (5.3).

Proof— Step 1: Let $X_{1;\alpha_1\alpha_2}(h_1)$ where $0 \leq h_1 \leq N_1$, and $\alpha_i = u$ (i.e., m_i is up) or d (i.e., m_i is down), $i = 1, 2$, be the stationary density of buffer occupancy. The densities $X_{1;\alpha_1\alpha_2}(h_1)$ and probabilities $Y_{1;\alpha_1\alpha_2}(0)$ and $Y_{1;\alpha_1\alpha_2}(N_1)$ are calculated in [9] and summarized in Table I.

Step 2: Using Table I, we show that

$$\begin{aligned}
 \text{PR} &= e_2 - Y_{1;du}(0) \\
 &= e_1 - Y_{1;ud}(N_1).
 \end{aligned}$$

Step 3: Using (5.3), (A.1), and the above expressions, we show that

$$\begin{aligned}
 \text{PR} &= e_2[1 - Q(p_1, r_1, p_2, r_2, N_1)] \\
 &= e_1[1 - Q(p_2, r_2, p_1, r_1, N_1)].
 \end{aligned}$$

Proof of Lemma 5.1: From Definition 5.1

$$\begin{aligned}
 mb_1 &= \text{Prob}(\{m_1 \text{ is up at time } t\} \\
 &\quad \cap \{b_1 \text{ is full at time } t\} \\
 &\quad \cap \{m_2 \text{ is down at time } t\}) \\
 &= Y_{1;ud}(N_1).
 \end{aligned}$$

From Lemma B.1, we have

$$mb_1 = e_1 Q(p_2, r_2, p_1, r_1, N_1).$$

According to Definition 5.2

$$\begin{aligned}
 ms_2 &= \text{Prob}(\{m_1 \text{ is down at time } t\} \\
 &\quad \cap \{b_1 \text{ is empty at time } t\} \\
 &\quad \cap \{m_2 \text{ is up at time } t\}) \\
 &= Y_{1;du}(0).
 \end{aligned}$$

Using Lemma B.1, we obtain

$$ms_2 = e_2 Q(p_1, r_1, p_2, r_2, N_1).$$

Proof of Lemma 5.2: Follows immediately from Lemma B.1, since $e_1 = e_2$ implies that $ms_2 = mb_1$. ■

REFERENCES

- [1] J. T. Lim, S. M. Meerkov, and F. Top, "Homogeneous, asymptotically reliable serial production lines: Theory and a case study," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 524–532, 1990.
- [2] D. Jacobs and S. M. Meerkov, "A system-theoretic property of serial production lines: Improvability," *Int. J. Syst. Sci.*, vol. 26, no. 4, pp. 755–785, 1995.
- [3] N. Viswanadham and Y. Narahari, *Performance Modeling of Automated Manufacturing Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [4] J. A. Buzacott and J. G. Shanthikumar, *Stochastic Models of Manufacturing Systems*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- [5] S. B. Gershwin, *Manufacturing Systems Engineering*. Englewood Cliffs, NJ: Prentice Hall, 1994.
- [6] *The Toyota Production System*. Tokyo, Japan: Toyota Motor Corp., 1996.
- [7] S. R. Lawrence and A. H. Buss, "Economic analysis of production bottlenecks," *Math. Prob. Eng.*, vol. 1, no. 4, pp. 341–363, 1995.
- [8] C.-T. Kuo, J.-T. Lim, and S. M. Meerkov, "Bottlenecks in serial production lines: A system-theoretic approach," *Math. Prob. Eng.*, vol. 2, no. 3, pp. 233–276, Apr. 1996.
- [9] D. Jacobs, "Improvability in production systems: Theory and case studies," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. Michigan, Ann Arbor, 1993, pp. 54–55.
- [10] S.-Y. Chiang, C.-T. Kuo, and S. M. Meerkov, "Bottlenecks in Markovian production lines: A systems approach," Control Group Rep. CGR-97-10, Univ. Michigan, Ann Arbor, Apr. 1997.