OPERATIONS RESEARCH CENTER

Working Paper

Modeling and Analysis of Markovian Continuous Flow Production Systems with a Finite Buffer: A General Methodology and Applications

by

B. Tan
S. B. Gershwin

OR 381-07 **March 2007**

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Modeling and Analysis of Markovian Continuous Flow Production Systems with a Finite Buffer: Methodology and Applications

Barış Tan
Graduate School of Business
Koç University
Rumeli Feneri Yolu, Sarıyer
Istanbul, Turkey
btan@ku.edu.tr

Stanley B. Gershwin
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139-4307
USA
gershwin@mit.edu

April 9, 2007

Abstract

Fluid flow models are used in performance evaluation of production, computer, and telecommunication systems. In order to develop a methodology to analyze general Markovian continuous material flow production systems with a finite buffer, a general single-buffer fluid flow system is modelled as a continuous time, continuous-discrete state space stochastic process and the steady state distribution is determined. Various performance measures such as the production rate and the expected buffer level are determined from the steady-state distributions. The flexibility of this methodology allows analysis of a wide range of models by specifying only the transition rates and the flow rates associated with the discrete states of each stage. Therefore the method is proposed as a tool for performance evaluation of general Markovian continuous flow systems with a finite buffer. The solution methodology is illustrated by analyzing a production system where each machine has multiple up and down states associated with their quality characteristics in detail. Then four different models: a model with multiple unreliable machines in parallel in each stage, a model with a merge-type structure, and a model with phase-type failure and repair-time distributions, and a model with multiple unreliable machines in series in each stage are analyzed by using the same methodology.

[1]The first version of this report was published on March 14, 2007

Contents

1 Introduction

In this study, we consider a two-stage continuous flow system separated by a finite capacity buffer (Figure 1). The dynamics of each stage is described by a continuous-time, discrete-state Markov chain where a different flow rate is associated with each state. This model can represent a wide range of systems. For example, it may represent a portion of a factory in which a stage represents an unreliable machine that may have any one of a wide variety of up- and down-time distributions; or a machine with variable quality; or multiple machines in series or parallel without intermediate buffers. For another example, it can represent a communications network in which message flow rates change according to Markov processes. In the following, we use the terms *stage* and *machine* interchangeably.

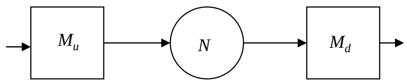

Figure 1: A Single Buffer Fluid Flow System with Two Stages

There is a vast literature on continuous material flow models of unreliable production lines. Two station single buffer unreliable continuous flow production systems have been analyzed in various studies, (e.g. Wijngaard 1979, Gershwin and Schick 1980, Dubois and Forestier 1982, Yeralan, Franck, and Quasem 1986, Yeralan and Tan 1997, among others). In most of these studies, each unreliable machine has two states: a single up state that represents the condition of a fully productive machine and a single down state that represent the condition where the machine is not productive due to a failure and the failure and repair times are exponential random variables.

More detailed models of production systems where each stage is modelled by using more than two states have been used to approximate general processing, failure, and repair time distributions by using phase-type distributions (Altıok 1985, Dallery 1994, Özdoğru and Altıok 2003); to study quality-quantity interactions (Tempelmeier and Burger 2001, Poffe and Gershwin 2005); or to develop new approximation methods with multiple up and down states (Levantesi, Matta, and Tolio 2003). Similarly, analysis of production lines with series or parallel structures (Mitra 1988, Patchong and Willaeys 2001), or merge structures (Tan 2001, Helber and Jusic 2004, Diamantidis, Papadopoulos, and Vidalis 2004) also received attention.

Although a variety of models are used to evaluate the performance of continuous flow production systems, currently there exists no unified methodology to analyze these systems. In the analysis of continuous flow models, once the state space is determined based on the underlying assumptions, the steady-state distribution is determined by analyzing the continuous time-continuous and discrete state space Markov process. In order to analyze this process, a set of differential equations that describe the behavior of the system is derived and then solved subject to boundary and normalization conditions. Without a general methodology, this process is repeated for each new model and considerable effort is required to model and to analyze any given system. This study is motivated by the need to develop a unified methodology to analyze all Markovian single-buffer continuous-flow production systems. With a similar objective, Gershwin and Fallah-Fini (2007) recently proposed a method to analyze general discrete-time, discrete-material-flow production lines with single buffer and identical processing rates. For a multiserver queue with Coxian arrival and service times and infinite waiting space, Bertsimas (1990) presents an algorithm to determine the system-size distribution and related performance measures. This approach can be used to analyze a special class of two-stage discrete-material, continuous-time production lines with an infinite buffer. In this line, the processing time of the first machine is a Coxian random variable. The second stage has a number of identical machines in parallel and the processing time of each machine is also a Coxian random variable.

Fluid flow models with a single buffer are also used to evaluate the performance of computer and telecommunication systems, (e.g. Anick, Mitra, and Sondhi 1982 and Elwalid and Mitra 1991). Recently, different methodologies are proposed to analyze general fluid flow models of computer and telecommunication systems with a finite buffer, (e.g. Serucola 2001, Ahn and Ramaswami 2003, Ahn, Jeon, and Ramaswami 2005, Soares and Latouche 2006). Although the fluid flow models developed for production and computer/telecommunication systems are similar, the methods developed for telecommunication and computer systems cannot be used to analyze production systems directly. The main difference between the models of telecommunication and computer systems and the models of production systems is the operation dependent failures that are observed in production systems. When the failures are operation dependent, an idle machine that is blocked or starved cannot fail. If a machine is partially blocked or partially starved and operating at a reduced rate, its failure rate will be lower than its rate when the buffer is partially full. As a result, the boundary processes when the buffer is empty or full are not the same as the interior process and all three processes must be analyzed accordingly.

In this paper, we present a methodology to analyze general Markovian continuous flow production systems with a finite buffer. The dynamics of the process when the buffer is partially full is determined by solving a set of first-order differential and algebraic equations. The unknown coefficients of the solution are determined by using a level crossing analysis. Namely, we first determine the probabilities of entering and exiting the full- and empty-buffer processes while the machines are in specific states by using a level crossing analysis. Then we link the entry and exit probabilities by using the conditional probabilities that are derived from the boundary processes. The only inputs of the model are the transition rates of each stage, the processing rates associated with the discrete states of each stage, and the buffer size. Therefore our model is quite general and allows analysis of a wide range of models by determining the required inputs. We illustrate our methodology by using a detailed example of a production system with multiple up and down states. We also discuss how different models can be analyzed by using our methodology.

The organization of the remaining part of the manuscript is as follows: In Section 2, we present a specific model where each stage has multiple up and down states corresponding to their quality characteristics to describe the types of models that can be analyzed with our methodology. In Section 3, we give a description of the general model, its assumptions, and introduce the variables used in the model. In Section 4, we present our methodology to analyze the general model and determine the performance measures of interest. In Section 5, the methodology is illustrated by analyzing the system described in Section 2. Then four different models: a model with multiple

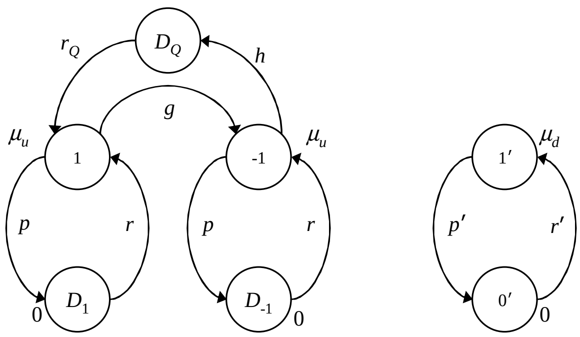

Figure 2: State transition diagram of the system of Section 2 and 5 with multiple up and down states

unreliable machines in parallel in each stage, a model with a merge-type structure, and a model with phase-type failure and repair-time distributions, and a model with multiple unreliable machines in series in each stage are analyzed by using the same methodology in Section 6 Finally, conclusions are given in Section 7.

2 Example

Before presenting the methodology to analyze general Markovian continuous flow systems, we first introduce a specific example to illustrate the type of models and also to show the generality of our methodology. In this example, we consider a production system with two unreliable machines with multiple up and down states and a finite buffer studied by Poffe and Gershwin (2005).

In the system we consider, the first stage has two up (State 1 and State -1) and three down states (State D_1, D_{-1}, and D_Q). We refer to states of M_u as down when the processing rate in that state is 0. In State 1, the machine produces products with no quality problems but when it is in State -1, the quality of the products produced is not perfect. Furthermore, the machine is subject to two different failures: operational failures (State D_1 and State D_{-1}) and quality failures (State D_Q) and they have different mean times to repair. Since these failures are different in nature, they cannot be modelled with a single down state. The failure rate is reduced proportionally when the processing rate of the machine is reduced due to starvation and blockage. The operational dependent failure mechanism is described in detail in Section 5. The second stage has one up (State $1'$) and one down state (State $0'$). Similar to the previous case, we refer to state $0'$ of the second stage as down because the processing rate of M_d is 0 in that state. More detailed representation of quality and unreliability characteristics of machines allows us to investigate quality and quantity issues jointly in the design and operation of production systems.

The processing rates of the upstream stage in both of the up states are equal to μ_u; the processing rate of the downstream stage in its up state is μ_d; and the processing rates of all the down states for both stages are equal to 0. Figure 2 depicts the state transitions for M_u and M_d for this model.

In order to analyze this system, a set of differential equations that describe the dynamics of the system must be derived and then solved subject to the boundary conditions. For example, when $\mu_u \neq \mu_d$, this model yields 7 first-order differential equations and 3 algebraic equations. In order to determine the steady-state probability distributions, 7 equations must be derived from the boundary processes. Once the steady-state distribution is obtained all the performance measures of interest can be determined from the distribution. Poffe and Gershwin (2005) derive these equations and solves them explicitly.

In order to analyze another system, e.g., an extension of this system where the second station is also modelled with 3 up and 2 down states, the same procedure must be repeated to derive all the equations (for this model when $\mu_u \neq \mu_d$, 16 differential and 9 algebraic equations) and then they must be solved.

In the next section, we present a methodology to analyze general Markovian continuous flow systems with a finite buffer. The specific model presented in this section is analyzed in Section 5 in order to explain the methodology in detail.

3 General Model

We consider a continuous flow system with two stages separated by a buffer with capacity N (Figure 1). The state of the system at time t is $s(t) = (X, \alpha_u, \alpha_d)$ where $0 \leq X \leq N$ is the buffer level, $\alpha_u \in \{1, ..., I_u\}$ is the state of the upstream stage M_u and $\alpha_d \in \{1, ..., I_d\}$ is the state of the downstream stage M_d. There are $I_u I_d$ discrete states in the state space $(\alpha_u, \alpha_d) \in S_M$. Figure 3 shows a sample realization of the system described in Section 2.

The maximum processing rate of M_u in state i is $\mu_i^u \geq 0$ and the maximum processing rate of M_d in state j is $\mu_j^d \geq 0$. The machines operate at their maximum rates unless they are starved or blocked. With these definitions, states need not be classified as up or down states as most of the other studies in the literature. A state with a maximum processing rate equal to zero can be considered as a down state.

When the buffer is empty in the machine state $(\alpha_u, \alpha_d) = (i, j)$ with $\mu_i^u = 0$ and $\mu_j^d > 0$ then M_d is said to be *completely starved* and it is forced to stop. However, when the buffer is empty and $\mu_j^d > \mu_i^u > 0$, M_d is said to be *partially starved* and it can continue its production at a reduced rate of μ_i^u. When the buffer is full in machine state $(\alpha_u, \alpha_d) = (i, j)$ with $\mu_i^u > 0$ and $\mu_j^d = 0$ then M_u is said to be *completely blocked* and the flow into the buffer is stopped. However, in the same state if $\mu_i^u > \mu_j^d > 0$, M_u is said to be *partially blocked* and it can continue its production at a reduced rate of μ_j^d. We assume that M_u is never starved and M_d is never blocked.

We partition the discrete states of the system into three sets depending on whether the buffer level goes up (Υ), down (Δ), or stays the same (Z) in that state

$$\begin{cases} (i,j) \in \Upsilon & \text{if} \quad \mu_i^u > \mu_j^d \\ (i,j) \in \Delta & \text{if} \quad \mu_i^u < \mu_j^d \\ (i,j) \in Z & \text{if} \quad \mu_i^u = \mu_j^d \end{cases}$$

and $S_M = \Upsilon \cup \Delta \cup Z$. The number of states in each of these sets are $I_\Upsilon = |\Upsilon|$, $I_\Delta = |\Delta|$, and $I_Z = |Z|$ respectively and $I_\Upsilon + I_\Delta + I_Z = I_u I_d$.

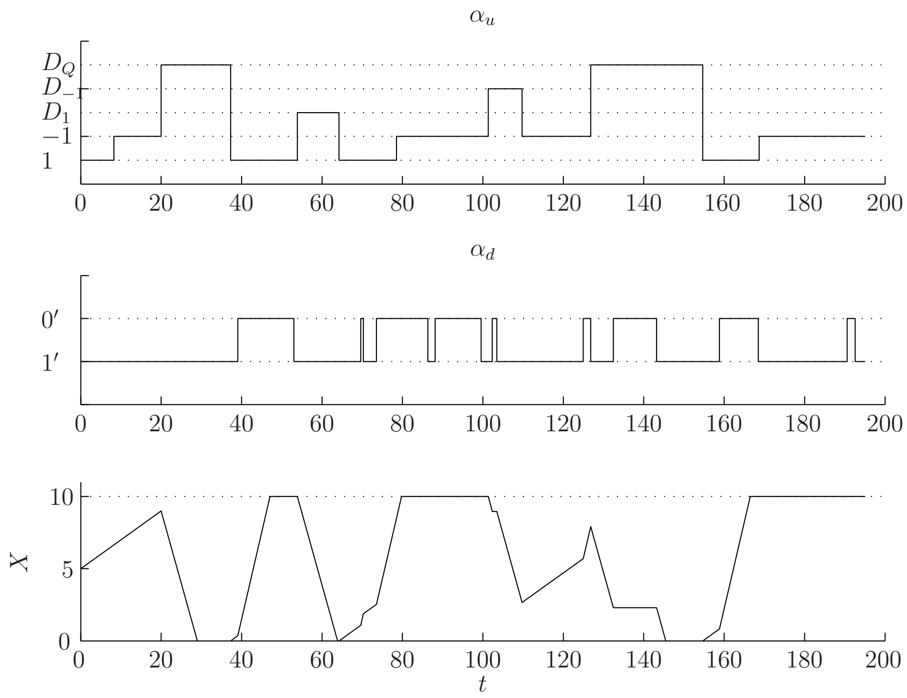

Figure 3: Sample path for the system of Section 2 and 5 with multiple up and down states ($\mu_u = 1.2$, $\mu_d = 1$, $p = 0.01$, $r = 0.1$, $p' = 0.05$, $r' = 0.10$, $g = 0.05$, $h = 0.10$, $r_Q = 0.10$)

For M_u, when $0 < X < N$, the transition time from state i to state i' is an exponential random variable with rate $\lambda^u_{ii'}$. Similarly for M_d, the transition time from state j to state j' is an exponential random variable with rate $\lambda^d_{jj'}$. When M_u is partially blocked, the transition time from state i to state i' is also an exponential random variable with rate $\psi^u_{ii'}$. Similarly, when M_d is partially starved, the transition rate from state j to state j' is $\psi^d_{jj'}$.

The time-dependent probability density while the buffer is partially full is

$$f(x,i,j,t) = \frac{\partial}{\partial x}\text{prob}[X(t) \le x, \alpha_u(t) = i, \alpha_d(t) = j] \;\text{ for }\; 0 < x < N.$$

We assume that the process is ergodic and the steady-state probabilities exist. The steady-state density functions are defined as

$$f(x,i,j) = \lim_{t\to\infty} f(x,i,j,t) \;\text{ for }\; 0 < x < N \tag{1}$$

and arranged in column vectors as

$$\mathbf{f}_S(x) = \{f(x,i,j)\}, \;\text{ for }\; (i,j) \in S, \;\; S = \Upsilon, \Delta, Z. \tag{2}$$

The probability of state $(0,i,j)$ at time t when the buffer is empty is denoted by $p(0,i,j,t)$ and the probability of state (N,i,j) at time t when the buffer is full is denoted by $p(N,i,j,t)$. The steady-state probabilities at the empty and full buffer states when $(\alpha_u, \alpha_d) = (i,j)$ are $p(0,i,j) = \lim_{t\to\infty} p(0,i,j,t)$ and $p(N,i,j) = \lim_{t\to\infty} p(N,i,j,t)$ respectively.

4 Analysis of Interior and Boundary Processes

In this section, the steady-state distribution is determined by analyzing the continuous time, continuous and discrete state space Markov process. First, the differential equations that describe the dynamics of the system when the buffer is in the interior $(0 < X < N)$ and when the buffer is at the boundary, i.e. when the buffer is empty $(X = 0)$ or full $(X = N)$, are derived. Then a solution technique is developed.

4.1 Interior Process

State Transition Equations Relating the probability density of the state at time $t + h$ to the probability density of the state at time t yields

$$f(x,i,j,t+h) = f(x - (\mu^u_i - \mu^d_j)h, i, j, t)\left(1 - \sum_{\substack{i'=1 \\ i' \ne i}}^{I_u} \lambda^u_{ii'} h\right)\left(1 - \sum_{\substack{j'=1 \\ j' \ne j}}^{I_d} \lambda^d_{jj'} h\right)$$

$$+ \sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} f(x-(\mu_{i'}^u - \mu_j^d)h, i', j, t)\lambda_{i'i}^u h \left(1 - \sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} \lambda_{jj'}^d h\right)$$

$$+ \sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} f(x-(\mu_i^u - \mu_{j'}^d)h, i, j', t)\lambda_{j'j}^d h \left(1 - \sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} \lambda_{ii'}^u h\right), \ (i,j) \in S_M. \quad (3)$$

The above equation can also be written in differential form by setting $h \to 0$ as

$$\frac{\partial f(x,i,j,t)}{\partial t} + (\mu_i^u - \mu_j^d)\frac{\partial f(x,i,j,t)}{\partial x} = -f(x,i,j,t)\left(\sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} \lambda_{ii'}^u + \sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} \lambda_{jj'}^d\right)$$

$$+ \sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} f(x,i',j,t)\lambda_{i'i}^u + \sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} f(x,i,j',t)\lambda_{j'j}^d, \ (i,j) \in S_M. \quad (4)$$

In steady state, the above equation yields $I_u I_d$ equations given below:

$$(\mu_i^u - \mu_j^d)\frac{\partial f(x,i,j)}{\partial x} = -f(x,i,j)\left(\sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} \lambda_{ii'}^u + \sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} \lambda_{jj'}^d\right)$$

$$+ \sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} f(x,i',j)\lambda_{i'i}^u + \sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} f(x,i,j')\lambda_{j'j}^d, \ (i,j) \in S_M. \quad (5)$$

Solution of the Internal Equations Note that the coefficient of $\frac{\partial f(x,i,j)}{\partial x}$ in Equation (5) can be positive, negative, or zero. Then the internal equations given in Equation (5) can be written in matrix form as

$$\begin{bmatrix} \frac{\partial \mathbf{f}_\Upsilon(x)}{\partial x} \\ \frac{\partial \mathbf{f}_\Delta(x)}{\partial x} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} \mathbf{f}_\Upsilon(x) \\ \mathbf{f}_\Delta(x) \\ \mathbf{f}_Z(x) \end{bmatrix} \quad (6)$$

where A_1 is a square matrix of size $(I_\Upsilon + I_\Delta) \times (I_\Upsilon + I_\Delta)$, A_4 is a square matrix of size $I_Z \times I_Z$, A_2 is a matrix of size $(I_\Upsilon + I_\Delta) \times I_Z$, A_3 is a matrix of size $I_Z \times (I_\Upsilon + I_\Delta)$, and $\mathbf{0}$ is a column vector of length I_Z. These matrices are determined by the parameters of the system.

Expanding Equation (6) gives the following set of equations:

$$\begin{bmatrix} \frac{\partial \mathbf{f}_\Upsilon(x)}{\partial x} \\ \frac{\partial \mathbf{f}_\Delta(x)}{\partial x} \end{bmatrix} = A_1 \begin{bmatrix} \mathbf{f}_\Upsilon(x) \\ \mathbf{f}_\Delta(x) \end{bmatrix} + A_2 \mathbf{f}_Z(x) \tag{7}$$

and

$$\mathbf{0} = A_3 \begin{bmatrix} \mathbf{f}_\Upsilon(x) \\ \mathbf{f}_\Delta(x) \end{bmatrix} + A_4 \mathbf{f}_Z(x). \tag{8}$$

Equation (8) is a set of algebraic equations and it can be solved directly to express $\mathbf{f}_Z(x)$ in terms of $\mathbf{f}_\Upsilon(x)$ and $\mathbf{f}_\Delta(x)$ as

$$\mathbf{f}_Z(x) = \Omega \begin{bmatrix} \mathbf{f}_\Upsilon(x) \\ \mathbf{f}_\Delta(x) \end{bmatrix} \tag{9}$$

where $\Omega = -A_4^{-1} A_3$. Since the square matrix A_4 has non-zero diagonal elements where each diagonal element is the sum of the transition rates from state $(i, j) \in Z$, A_4^{-1} always exists.

Now, inserting Equation (9) into (8) yields a first-order matrix differential equation given below

$$\begin{bmatrix} \frac{\partial \mathbf{f}_\Upsilon(x)}{\partial x} \\ \frac{\partial \mathbf{f}_\Delta(x)}{\partial x} \end{bmatrix} = \Lambda \begin{bmatrix} \mathbf{f}_\Upsilon(x) \\ \mathbf{f}_\Delta(x) \end{bmatrix}. \tag{10}$$

where $\Lambda = A_1 - A_2 A_4^{-1} A_3$. The solution of this first-order matrix differential equation is

$$\begin{bmatrix} \mathbf{f}_\Upsilon(x) \\ \mathbf{f}_\Delta(x) \end{bmatrix} = e^{\Lambda x} \mathbf{w} \tag{11}$$

where $e^{\Lambda x}$ is a matrix exponential determined by matrix Λ and $\mathbf{w}$ is a column vector of length $I_\Upsilon + I_\Delta$.

Inserting the solution for $\mathbf{f}_\Upsilon(x)$ and $\mathbf{f}_\Delta(x)$ given in Equation (11) into Equation (9) yields the solution for $\mathbf{f}_Z(x)$:

$$\mathbf{f}_Z(x) = \Omega e^{\Lambda x} \mathbf{w}. \tag{12}$$

When vector $\mathbf{w}$ is determined, all the density functions are determined by Equations (11) and (12). Since the length of $\mathbf{w}$ is $I_\Upsilon + I_\Delta$, $I_\Upsilon + I_\Delta$ equations are needed to determine the weights uniquely. We determine these equations by analyzing the boundary processes in the following. We first discuss important concepts and results of level crossing analysis.

4.2 Level Crossing

In order to relate the densities of the partially-full buffer process and the boundary buffer processes when the buffer is empty or full, we use a level crossing analysis similar to the one utilized in Yeralan and Tan (1997). With this approach, the entry and exit probabilities into the empty- and full-buffer processes are determined from the density functions.

In order to explain this approach, first note that since the buffer is finite and the process is ergodic, any process realization must visit any given buffer level an infinite times in the long run.

Equivalently, at a buffer level x, the number of upward crossings in a given time period is equal to the number of downward crossings in the same time period in the long run.

In order to define level crossing at the buffer level x formally, consider the following event

$$G(x,i,j,h) = \{\alpha_u = i,\ \alpha_d = j,\ 0 \leq x \leq X \leq x+h \leq N\}.$$

A level crossing at x when $(\alpha_u, \alpha_d) = (i,j)$ is defined as event $G(x,i,j,h)$ as $h \to 0$ when $(i,j) \in \Upsilon \cup \Delta$.

The rate of change of buffer level in state (i,j) is $|\mu_i^u - \mu_j^d|$. Note that $|\mu_i^u - \mu_j^d| \neq 0$ only when $(i,j) \in \Upsilon \cup \Delta$. Once event $G(x,i,j,h)$ occurs, it will last for $\frac{h}{|\mu_i^u - \mu_j^d|}$ time units. Therefore when $h \to 0$, the probability that (α_u, α_d) changes its state in this time period goes to zero.

The probability of this event can be determined as

$$\text{prob}[G(x,i,j,h)] = \int_x^{x+h} f(x',i,j)dx' = f(x,i,j)h + o(h). \tag{13}$$

Let $L(x,i,j,T)$ denote the number of level crossings in state (x,i,j) in the time interval $[t, t+T]$. If the system is observed for T units of time as $T \to \infty$, the probability that event $G(x,i,j,h)$ occurs is also the fraction of time the system spends in set $G(x,i,j,h)$ in the long run. Consequently,

$$f(x,i,j)h = \lim_{T\to\infty} \frac{L(x,i,j,T)}{T} \frac{h}{|\mu_i^u - \mu_j^d|} + o(h). \tag{14}$$

Therefore, the expected number of level crossings per unit time in the long run is determined by the densities and the flow rates as

$$\lim_{T\to\infty} \frac{L(x,i,j,T)}{T} = |\mu_i^u - \mu_j^d| f(x,i,j). \tag{15}$$

In state (i,j) with $\mu_i^u > \mu_j^d$, $(\mu_i^u - \mu_j^d)f(x,i,j)$ is the expected number of upward crossings at buffer level $0 < x < N$ per unit time. Similarly in state (i,j) with $\mu_i^u < \mu_j^d$, $(\mu_j^d - \mu_i^d)f(x,i,j)$ is the expected number of downward crossings per unit time. Since at any given buffer level, the expected number of upward and downward crossings are equal in the long run, we can also write

$$\sum_{i=1}^{I_u} \sum_{j=1}^{I_d} (\mu_i^u - \mu_j^d) f(x,i,j) = 0. \tag{16}$$

The above result can also be derived by adding all the transition equations given in Equation (5) for $(i,j) \in S_M$ that yields the level crossing equivalence in differential form. For other results in level crossing analysis, the reader is referred to Blake and Lindsey (1973) and Brill (1978).

In order to complete the analysis, we must determine $p(0,i,j)$, the steady-state probability that the buffer is empty and the machine states are $(\alpha_u, \alpha_d) = (i,j)$; and $p(N,i,j)$, the steady-state probability that the buffer is full and the machine states are $(\alpha_u, \alpha_d) = (i,j)$. Next we focus on the $x = 0$ boundary.

4.3 Empty Buffer Process

Now we derive the equations that describe the dynamics of the system when the buffer is empty, or becomes empty, or stops being empty. As the buffer level decreases in states $(i,j) \in \Delta$, the buffer eventually becomes empty if no other transition occurs first. Once the buffer becomes empty, it stays empty until the system makes a transition to a state $(i,j) \in \Upsilon$. When the buffer is empty, the set of states where the buffer stays empty is $S_0 = \Delta \cup Z$ and $I_{S_0} = |S_0| = I_\Delta + I_Z$.

State Transition Equations Let t_k^0 be the kth time the buffer becomes empty. We calculate $p(0,i,j)$ from the state transition rates that determine $\pi(0,i,j,t_k^0+\tau)$, the probability that $X=0$ and $(\alpha_u,\alpha_d)=(i,j)$ at time $t_k^0+\tau$ given that the buffer became empty at time t_k^0 and has been empty during $[t_k^0, t_k^0+\tau]$.

In order for the buffer to become empty at time t_k^0 the machine state (α_u,α_d) must have been in set Δ at that time. For it to stay empty during $[t_k^0, t_k^0+\tau]$, (α_u,α_d) must be in set $S_0 = \Delta \cup Z$ during that interval. For it to become non-empty at time $t_k^0+\tau$, (α_u,α_d) must make a transition into Υ at that time.

The dynamics of the system during an interval when the buffer stays empty are given by the following equations:

$$\begin{aligned} \frac{d\pi(0,i,j,\tau)}{d\tau} &= -\pi(0,i,j,\tau)\left(\sum_{\substack{i'=1 \\ i'\neq i}}^{I_u} \lambda_{ii'}^u + \sum_{\substack{j'=1 \\ j'\neq j}}^{I_d} \psi_{jj'}^d \right) \\ &+ \sum_{\substack{i'=1 \\ i'\neq i \\ (i',j)\in S_0}}^{I_u} \pi(0,i',j,\tau)\lambda_{i'i}^u + \sum_{\substack{j'=1 \\ j'\neq j \\ (i,j')\in S_0}}^{I_d} \pi(0,i,j',\tau)\psi_{j'j}^d, \quad (i,j)\in S_0. \end{aligned} \tag{17}$$

Equation (17) can be written in matrix form as

$$\frac{d\pi_{S_0}^0(\tau)}{d\tau} = A_0 \pi_{S_0}^0(\tau) \tag{18}$$

where $\pi_{S_0}^0(\tau) = \{\pi(0,i,j,\tau)\}$ for $(i,j)\in S_0$ and A_0 is a $I_{S_0}\times I_{S_0}$ square matrix.

The empty buffer process ends with a transition into a state where the buffer level starts increasing. Let $q(0,i,j,t_k^0+\tau)$ be the rate at which the process enters into the state $(i,j)\in\Upsilon$ at time $t_k^0+\tau$ given that the buffer became empty at time t_k^0. This rate can be determined as

$$q(0,i,j,\tau) = \sum_{\substack{i'=1 \\ i'\neq i \\ (i',j)\in S_0}}^{I_u} \pi(0,i',j,\tau)\lambda_{i'i}^u + \sum_{\substack{j'=1 \\ j'\neq j \\ (i,j')\in S_0}}^{I_d} \pi(0,i,j',\tau)\psi_{j'j}^d, \quad (i,j)\in\Upsilon. \tag{19}$$

Equation (19) can be written in matrix form as

$$\mathbf{q}^0_\Upsilon(\tau) = B_0 \pi^0_{S_0}(\tau) \tag{20}$$

where $\mathbf{q}^0_\Upsilon(\tau) = \{q(0,i,j,\tau)\}$ for $(i,j) \in \Upsilon$ and B_0 is a $I_\Upsilon \times I_{S_0}$ matrix.

Entry and Exit Probabilities In order to link the interior and the empty buffer process, we first analyze how the buffer becomes empty and then how it exits the empty buffer states. Let us define a discrete time random process $\{\phi_k^{\texttt{enter}}, k = 1, 2, ...\}$ sampled from the process $\{s(t), t \geq 0\}$ at the instances t_k^0, $k = 1, 2, ...$ when the buffer becomes empty. The random variable $\phi_k^{\texttt{enter}}$ consists of states of the machines at the instant when the buffer becomes empty for the kth time. That is, if $X(t_k^0 - h) = 0^+$ and $X(t_k^0) = 0$ as $h \to 0$ then $\phi_k^{\texttt{enter}} = (\alpha_u(t_k^0), \alpha_d(t_k^0))$. The subscript k is dropped to represent $\phi_k^{\texttt{enter}}$ in steady state.

The probability that the buffer becomes empty while the machines are in state (i,j) is the ratio of the number of downward crossings in this particular state to the number of all possible downward crossings at $X = 0^+$:

$$\begin{aligned} \text{prob}[\phi^{\texttt{enter}} = (i,j)] &= \lim_{T\to\infty} \frac{L(0^+, i, j, T)/T}{\sum\limits_{(i',j')\in\Delta} L(0^+, i', j', T)/T} \\ &= \frac{(\mu_i^u - \mu_j^d) f(0^+, i, j)}{\sum\limits_{(i',j')\in\Delta} (\mu_{i'}^u - \mu_{j'}^d) f(0^+, i', j')}, \quad (i,j) \in \Delta. \end{aligned} \tag{21}$$

Similarly, let us define another discrete time random process $\{\phi_k^{\texttt{exit}}, k = 1, 2, ...\}$ sampled from the process $\{s(t), t \geq 0\}$ at the instances τ_k^0, $k = 1, 2, ...$ when the buffer level starts increasing following being empty. The random variable $\phi_k^{\texttt{exit}}$ describes the states of the machines at the instant when the buffer level starts increasing after being empty for the kth time. That is, if $X(\tau_k^0 - h) = 0$ and $X(\tau_k^0) = 0^+$ as as $h \to 0$ then $\phi_k^{\texttt{exit}} = (\alpha_u(\tau_k^0), \alpha_d(\tau_k^0))$. Then the probability that the process exits the empty buffer state with a transition into state $(i,j) \in \Upsilon$ is given as

$$\text{prob}[\phi^{\texttt{exit}} = (i,j)] = \frac{(\mu_i^u - \mu_j^d) f(0^+, i, j)}{\sum\limits_{(i',j')\in\Upsilon} (\mu_{i'}^u - \mu_{j'}^d) f(0^+, i', j')}, \quad (i,j) \in \Upsilon. \tag{22}$$

The empty buffer process relates the probabilities given in Equations (21) and (22). More specifically,

$$\text{prob}[\phi^{\texttt{exit}} = (i,j)] = \sum_{(i',j')\in\Delta} \text{prob}[\phi^{\texttt{exit}} = (i,j) \mid \phi^{\texttt{enter}} = (i',j')]\text{prob}[\phi^{\texttt{enter}} = (i',j')], \quad (i,j) \in \Upsilon. \tag{23}$$

Inserting Equations (21) and (22) into Equation (23) and using the equivalence of the upward and downward crossings given in Equation (16) yields

$$(\mu_i^u - \mu_j^d) f(0^+, i, j) = \sum_{(i',j') \in \Delta} \text{prob}[\phi^{\texttt{exit}} = (i, j) \mid \phi^{\texttt{enter}} = (i', j')](\mu_{j'}^d - \mu_{i'}^u) f(0^+, i', j'),$$
$$(i, j) \in \Upsilon \qquad (24)$$

where conditional probabilities $\text{prob}[\phi^{\texttt{exit}} = (i, j) \mid \phi^{\texttt{enter}} = (i', j')]$ are determined from Equations (18) and (20).

The $(i, j)(i', j')$ element of $-B_0 A_0^{-1}$ is the conditional probability that the empty buffer process exits in a particular state $(i, j) \in \Upsilon$ given that it starts in one of the states $(i', j') \in S_0$ where the buffer stays empty. Since the empty buffer process can start only in states $(i, j) \in \Delta$, let G_0 be a $I_\Upsilon \times I_\Delta$ matrix that is obtained by eliminating the columns of $-B_0 A_0^{-1}$ corresponding to states $S_0 \setminus \Delta$. Accordingly, by using the solution of the density functions given in Equation (11), Equation (24) can be written in matrix form as

$$\left[\text{diag}(\mathbf{m}_\Upsilon) \;\; 0_{I_\Upsilon \times I_\Delta} \right] \mathbf{w} = G_0 \left[0_{I_\Delta \times I_\Upsilon} \;\; \text{diag}(\mathbf{m}_\Delta) \right] \mathbf{w} \qquad (25)$$

where $\mathbf{m}_\Upsilon = \{(\mu_i^u - \mu_j^d) | \; (i, j) \in \Upsilon\}$ and $\mathbf{m}_\Delta = \{(\mu_j^d - \mu_i^u) | \; (i, j) \in \Delta\}$. We use $\text{diag}(\mathbf{a})$ to represent a diagonal matrix formed with the elements of vector $\mathbf{a}$ and $0_{k \times l}$ is a $k \times l$ matrix of zeros.

Since $\sum_{(i,j) \in \Upsilon} \text{prob}[\phi^{\texttt{exit}} = (i, j)] = 1$, Equation (25) gives $I_\Upsilon - 1$ linearly independent equations that will be used to determine $\mathbf{w}$.

Steady-State Probability Distribution Due to ergodicity of the process, the probability that $X = 0$ and $(\alpha_u, \alpha_d) = (i, j)$ is also the fraction of the total time the process stays in this state in a given time period the long run.

We can determine the total time the process stays in state $(i, j) \in S_0$ while $X = 0$ in a given time period by determining the number of times the buffer becomes empty and the time the process stays in this state for each time the buffer becomes empty in the same time period.

Given that the machine states $(\alpha_u, \alpha_d) = (i', j') \in \Delta$ at the time the buffer becomes empty, the expected time that the machine states (α_u, α_d) stay in $(i, j) \in S_0$ before exiting to a state $(\alpha_u, \alpha_d) \in \Upsilon$ is denoted by $E[T^0_{(i,j),(i',j')}]$. Then, the steady-state probability of state $(0, i, j)$, $p(0, i, j)$ is given as

$$p(0, i, j) = \sum_{(i',j') \in \Delta} \lim_{T \to \infty} \frac{L(0^+, i', j', T) E[T^0_{(i,j),(i',j')}]}{T} \qquad (26)$$

By using Equation (15), we can write $p(0, i, j)$ in terms of the densities, processing rates, and expected sojourn times as

$$p(0, i, j) = \sum_{(i',j') \in \Delta} (\mu_{j'}^d - \mu_{i'}^u) f(0^+, i', j') E[T^0_{(i,j),(i',j')}]. \qquad (27)$$

The $(i, j), (i', j')$ element of matrix $-A_0^{-1}$ determined from Equation (18) gives the expected sojourn time in state $(i, j) \in S_0$ given that (α_u, α_d) starts in state $(i', j') \in S_0$. Since the empty buffer process can start only in states $(i, j) \in \Delta$, we define $E[T^0]$ to be an $I_{S_0} \times I_\Delta$ matrix that is obtained by eliminating the columns of $-A_0^{-1}$ corresponding to states in S_0 that are not in Δ, i.e., $S_0 \setminus \Delta$.

Using the solution of the density functions given in Equation (11) in Equation (27) gives

$$\mathbf{p}_0 = E[T^0] \left[\begin{array}{cc} 0_{I_\Delta \times I_\Upsilon} & \text{diag}(\mathbf{m}_\Delta) \end{array} \right] \mathbf{w} \tag{28}$$

where $\mathbf{p}_0 = \{p(0, i, j)\}$.

Now, we can also determine the probability that the buffer is empty as

$$\text{prob}[X = 0] = \sum_{(i',j') \in S_0} p(0, i, j) = u_{I_{S_0}} \mathbf{p}_0 \tag{29}$$

where $u_k = (1, 1, ..., 1)$ is a row vector of ones of length k.

4.4 Full Buffer Process

The last step is the analysis of the full buffer process. As the buffer level increases in states $(i, j) \in \Upsilon$, the buffer eventually becomes full if no other transition occurs first. Once the buffer becomes full, it stays full until the system makes a transition to a state $(i, j) \in \Delta$. When the buffer is full, the set of states where the buffer stays full is $S_N = \Upsilon \cup Z$ and $I_{S_N} = |S_N|$.

State Transition Equations Let t_k^N be the kth time the buffer becomes full. We calculate $p(N, i, j)$ from the state transition rates that define $\pi(N, i, j, t_k^N + \tau)$, the probability that $X = N$ and $(\alpha_u, \alpha_d) = (i, j)$ at time $t_k^N + \tau$ given that the buffer became full at time t_k^N and has been full during $[t_k^N, t_k^N + \tau]$.

In order for the buffer to become full at time t_k^N the machine state (α_u, α_d) must have been in set Υ at that time. For it to stay full during $[t_k^N, t_k^N + \tau]$, (α_u, α_d) must be in set $S_N = \Upsilon \cup Z$ during that interval. For it to become non-full at time $t_k^N + \tau$, (α_u, α_d) must make a transition into Δ at that time.

The dynamics of the system when the buffer stays full in state $(i, j) \in S_N$ are given as

$$\begin{aligned} \frac{d\pi(N, i, j, \tau)}{d\tau} &= -\pi(N, i, j, \tau) \left(\sum_{\substack{j'=1 \\ j' \neq j}}^{I_d} \lambda_{jj'}^d + \sum_{\substack{i'=1 \\ i' \neq i}}^{I_u} \psi_{ii'}^u \right) \\ &+ \sum_{\substack{j'=1 \\ j' \neq j \\ (i,j') \in S_N}}^{I_d} \pi(N, i, j', \tau) \lambda_{j'j}^d + \sum_{\substack{i'=1 \\ i' \neq i \\ (i',j) \in S_N}}^{I_u} \pi(N, i', j, \tau) \psi_{i'i}^u, \quad (i, j) \in S_N. \end{aligned} \tag{30}$$

The above equation can be written in matrix form as

$$\frac{\pi_{S_N}^N(\tau)}{d\tau} = A_N \pi_{S_N}^N(\tau) \tag{31}$$

where $\pi^N_{S_N}(\tau) = \{\pi(N,i,j,\tau)\}$ for $(i,j) \in S_N$ and A_N is a $I_{S_N} \times I_{S_N}$ square matrix

The full buffer process ends with a transition into a state where the buffer level starts decreasing. Let $q(N,i,j,t^N_k+\tau)$ be the rate at which the process enters into the state $(i,j) \in \Delta$ at time $t^N_k+\tau$ given that the buffer became full at time t^N_k. We can determine this rate as

$$q(N,i,j,\tau) = \sum_{\substack{j'=1 \\ j' \neq j \\ (i,j') \in S_N}}^{I_d} \pi(N,i,j',\tau)\lambda^d_{j'j} + \sum_{\substack{i'=1 \\ i' \neq i \\ (i',j) \in S_N}}^{I_u} \pi(N,i',j,\tau)\psi^u_{i'i}, \quad (i,j) \in \Delta \tag{32}$$

or in matrix form

$$\mathbf{q}^N_\Delta(\tau) = B_N \pi^N_{S_N}(\tau) \tag{33}$$

where $\mathbf{q}^N_\Delta(\tau) = \{q(N,i,j,\tau)\}$ for $(i,j) \in \Delta$ and B_N is a $I_\Delta \times I_{S_N}$ matrix.

Entry and Exit Probabilities In order to link the interior and the full-buffer process, we first analyze how the buffer becomes full and then how it exits the full buffer states and the buffer level starts decreasing. Let us define a discrete time random process $\{\varphi^{\texttt{enter}}_k, k = 1,2,...\}$ sampled from the process $\{s(t), t \geq 0\}$ at the instances where the buffer becomes full. The random variable $\varphi^{\texttt{enter}}_k$ describes the states of the machines at the instances t^N_k, $k = 1,2,...$ when the buffer becomes full for the kth time. That is, if $X(t^N_k - h) = N^-$ and $X(t^N_k) = N$ as $h \to 0$ then $\varphi^{\texttt{enter}}_k = (\alpha_u(t^N_k), \alpha_d(t^N_k))$. The subscript k is dropped to represent the random variable in steady state.

The probability that the buffer becomes full while the process has been in a specific state is the ratio of the number of upward crossings in this particular state and the all possible upward crossings at $X = N^-$:

$$\text{prob}[\varphi^{\texttt{enter}} = (i,j)] = \frac{(\mu^u_i - \mu^d_j)f(N^-,i,j)}{\sum_{(i',j') \in \Upsilon} (\mu^u_{i'} - \mu^d_{j'})f(N^-,i',j')}, \quad (i,j) \in \Upsilon. \tag{34}$$

Similarly, let us define another discrete time random process $\{\varphi^{\texttt{exit}}_k, k = 1,2,...\}$ sampled from the process $\{s(t), t \geq 0\}$ at the instances τ^N_k, $k = 1,2,...$ when the buffer level starts decreasing following being full. The random variable $\varphi^{\texttt{exit}}_k$ describes the states of the machines at the instant when the buffer level starts decreasing following being full for the kth time. That is, if $X(\tau^N_k - h) = N$ and $X(\tau^N_k) = N^-$ as $h \to 0$ then $\varphi^{\texttt{exit}}_k = (\alpha_u(\tau^N_k), \alpha_d(\tau^N_k))$. Then the probability that the process exits the full buffer state with a transition into state $(i,j) \in \Delta$ is given as

$$\text{prob}[\varphi^{\texttt{exit}} = (i,j)] = \frac{(\mu^u_i - \mu^d_j)f(N^-,i,j)}{\sum_{(i',j') \in \Delta} (\mu^u_{i'} - \mu^d_{j'})f(N^-,i',j')}, \quad (i,j) \in \Delta. \tag{35}$$

The full buffer process relates the probabilities given in Equations (34) and (35). More specifically,

$$\text{prob}[\varphi^{\texttt{exit}} = (i,j)] = \sum_{(i',j') \in \Upsilon} \text{prob}[\varphi^{\texttt{exit}} = (i,j) \mid \varphi^{\texttt{enter}} = (i',j')]\text{prob}[\varphi^{\texttt{enter}} = (i',j')],$$

$$(i,j) \in \Delta. \qquad (36)$$

Inserting Equations (34) and (35) into Equation (36) and simplifying by using the equivalence of the upward and downward crossings given in Equation (16) gives

$$(\mu_j^d - \mu_i^u) f(N^-, i, j) = \sum_{(i',j') \in \Upsilon} \text{prob}[\varphi^{\texttt{exit}} = (i,j) \mid \varphi^{\texttt{enter}} = (i',j')](\mu_{i'}^u - \mu_{j'}^d) f(N^-, i', j'),$$
$$(i,j) \in \Delta \qquad (37)$$

where conditional probabilities $\text{prob}[\varphi^{\texttt{exit}} = (i,j) \mid \varphi^{\texttt{enter}} = (i',j')]$ are determined from Equations (31) and (33).

More specifically, the $(i,j)(i',j')$ element of matrix $-B_N A_N^{-1}$ is the conditional probability that the full buffer process exits in a particular state $(i,j) \in \Delta$ given that it starts in one of the states $(i',j') \in S_N$ where the buffer stays full. Since the full buffer process can start only in states $(i,j) \in \Upsilon$, let G_N be a $I_\Delta \times I_\Upsilon$ matrix that is obtained by eliminating the columns of $-B_N A_N^{-1}$ corresponding to states $S_N \setminus \Upsilon$.

Using the solution of the density functions given in Equation (11) in Equation (37) yields

$$\left[\begin{array}{cc} 0_{I_\Delta \times I_\Upsilon} & \text{diag}(\mathbf{m}_\Delta) \end{array}\right] e^{\Lambda N} \mathbf{w} = G_N \left[\begin{array}{cc} \text{diag}(\mathbf{m}_\Upsilon) & 0_{I_\Upsilon \times I_\Delta} \end{array}\right] e^{\Lambda N} \mathbf{w}. \qquad (38)$$

Since $\sum_{(i,j) \in \Delta} \text{prob}[\varphi^{\texttt{exit}} = (i,j)] = 1$, Equation (38) gives $I_\Delta - 1$ linearly independent equations that will be used to determine $\mathbf{w}$.

Steady-State Distribution Ergodicity of the process ensures that the probability that $X = N$ and $(\alpha_u, \alpha_d) = (i,j)$ is also the ratio of the total time the process stays in this state in a given time period the long run.

The total time the process stays in state $(i,j) \in S_N$ while $X = N$ in a given time period can be determined by multiplying the number of times the buffer becomes full and the time the process stays in this state once the buffer becomes full in the same time period.

Given that the machine states $(\alpha_u, \alpha_d) = (i',j') \in \Upsilon$ at the time the buffer becomes full, the expected time that the machine states (α_u, α_d) stay in $(i,j) \in S_N$ before exiting to a state $(\alpha_u, \alpha_d) \in \Delta$ is denoted by $E[T^N_{(i,j),(i',j')}]$. Then, the steady-state probability of state (N, i, j), $p(N, i, j)$ is given as

$$p(N, i, j) = \sum_{(i',j') \in \Upsilon} \lim_{T \to \infty} \frac{L(N^-, i', j', T) E[T^N_{(i,j),(i',j')}]}{T} \qquad (39)$$

By using Equation (15), we can write $p(N, i, j)$ in terms of the densities, processing rates, and expected sojourn times as

$$p(N, i, j) = \sum_{(i',j') \in \Upsilon} (\mu_{i'}^u - \mu_{j'}^d) f(N^-, i', j') E[T^N_{(i,j),(i',j')}] \qquad (40)$$

The $(i,j),(i',j')$ element of matrix $-A_N^{-1}$ determined from Equation (31) gives the expected sojourn time in state $(i,j) \in S_N$ given that (α_u, α_d) starts in $(i',j') \in S_N$. Since the full buffer process can start only in states $(i,j) \in \Upsilon$, we define $E[T^N]$ to be an $I_{S_N} \times I_\Upsilon$ matrix that is obtained by eliminating the columns of $-A_N^{-1}$ corresponding to states in S_N that are not in Υ, i.e., $S_N \setminus \Upsilon$.

Inserting the solution of the density functions given in Equation (11) in Equation (40) yields

$$\mathbf{p}_N = E[T^N] \left[\begin{array}{cc} \text{diag}(\mathbf{m}_\Upsilon) & 0_{I_\Upsilon \times I_\Delta} \end{array} \right] e^{\Lambda N} \mathbf{w}. \tag{41}$$

where $\mathbf{p}_N = \{p(N,i,j)\}$.

Then the probability that the buffer is full can be calculated as

$$\text{prob}[X = N] = \sum_{(i,j)\in S_N} p(N,i,j) = u_{I_{S_N}} \mathbf{p}_N. \tag{42}$$

4.5 Solution of the Probability Densities

Once the weight vector $\mathbf{w}$ is determined, all the steady-state probabilities are also determined. Since there are $I_\Upsilon + I_\Delta$ weights and Equations (25) and (38) give a total of $I_\Upsilon + I_\Delta - 2$ equations, two additional equations are required to uniquely determine $\mathbf{w}$.

The first equation is the equivalence of the total upward and downward crossings in the interior region. Integrating Equation (16) from 0 to N yields

$$\int_0^N \sum_{i=1}^{I_u} \sum_{j=1}^{I_d} \mu_i^u f(x,i,j)dx = \int_0^N \sum_{i=1}^{I_u} \sum_{j=1}^{I_d} \mu_j^d f(x,i,j)dx \tag{43}$$

or in matrix form

$$\left[\begin{array}{cc} \mathbf{m}_\Upsilon & -\mathbf{m}_\Delta \end{array} \right] \left(\int_0^N e^{\Lambda x} dx \right) \mathbf{w} = 0. \tag{44}$$

The second equation is the normalization equation:

$$\sum_{i=1}^{I_u} \sum_{j=1}^{I_d} (p(0,i,j) + p(N,i,j)) + \int_0^N \sum_{i=1}^{I_u} \sum_{j=1}^{I_d} f(x,i,j)dx = 1. \tag{45}$$

By using Equations (11), (12), (29) and (42), the normalization equation can be written in matrix form as

$$\left(u_{I_{S_0}} E[T^0] \left[\begin{array}{cc} 0_{I_\Delta \times I_\Upsilon} & \text{diag}(\mathbf{m}_\Delta) \end{array} \right] + \nu \left(\int_0^N e^{\Lambda x} dx \right) + u_{I_{S_N}} E[T^N] \left[\begin{array}{cc} \text{diag}(\mathbf{m}_\Upsilon) & 0_{I_\Upsilon \times I_\Delta} \end{array} \right] e^{\Lambda N} \right) \mathbf{w} = 1 \tag{46}$$

where $\nu = (u_{I_\Upsilon + I_\Delta} + u_{I_Z} \Omega)$.

Now Equations (25) and (38) with Equations (44) and (46) give $I_\Upsilon + I_\Delta$ linearly independent equations that uniquely determine $\mathbf{w}$. Therefore all the steady-state probability distributions that describe the dynamics of the system are determined by these equations.

4.6 Performance Measures

When the probability densities are determined, all performance measures of interest can be calculated. In a production setting, the main performance measures of interest are the production rate and the expected buffer level.

The production rate is the amount of material processed per unit time in the long run. The production rate of the first stage and the second stage are the same due to the conservation of flow. Therefore we give the production rate of the first stage without loss of generality. The production rate in the internal states can be determined in a straight-forward way. Since the first stage can be forced to produce at a reduced rate due to partial blocking and the second stage can be forced to produce at a reduced rate due to partial starvation, this must be taken into consideration. The following equation gives the production rate of the first stage:

$$\Pi = \sum_{(i,j)\in S_0} \mu_i^u p(0,i,j) + \sum_{(i,j)\in S_M} \int_0^N \mu_i^u f(x,i,j)dx + \sum_{(i,j)\in S_N} \mu_j^d p(N,i,j) \tag{47}$$

The expected buffer level is determined as

$$E[X] = \sum_{i=1}^{I_u}\sum_{j=1}^{I_d}\left(\int_0^N xf(x,i,j)dx + Np(N,i,j)\right). \tag{48}$$

Once the steady-state distribution is determined, other performance measures of interest can also be evaluated directly.

5 Analysis of the Example

In this section, we analyze the specific system described in Section 2 by using our methodology. In order to explain the methodology, all the variables defined in Section 3 are given explicitly for this model. We also evaluate the performance of the system as some of the system parameters change.

Before defining the variables, we first discuss modelling operation dependent failures in this setting. Although our methodology is developed to work with arbitrary values of $\psi_{ii'}^u$ and $\psi_{jj'}^d$, in this example and in the production examples analyzed in Tan and Gershwin (2007), a specific case where the reduction in the transition rates at the boundaries is proportional to the reduction in the processing rate is considered similar to other papers in the literature (e.g. Gershwin and Schick 1980). That is when the buffer is empty and M_d is producing at a reduced rate of μ_i^u, $\psi_{jj'}^d = \frac{\mu_i^u}{\mu_j^d}\lambda_{jj'}^d$. This setting implies that when $\mu_i^u = 0$, $\psi_{jj'}^d = 0$ and therefore it is not possible to make a transition when M_d is completely starved. Similarly, when the buffer is full and M_u is producing at a reduced rate of μ_j^d, $\psi_{ii'}^u = \frac{\mu_j^d}{\mu_i^u}\lambda_{ii'}^d$. Similar to the previous case, when $\mu_j^d = 0$, $\psi_{ii'}^u = 0$ and therefore a transition is not possible when M_u is completely blocked.

5.1 Model Inputs

Our solution methodology requires only matrices $\lambda^u = \{\lambda_{ii'}^u\}$, $\lambda^d = \{\lambda_{jj'}^d\}$, $\psi^u = \{\psi_{ii'}^u\}$, $\psi^d = \{\psi_{jj'}^d\}$, vectors $\mu^u = \{\mu_i^u\}$, $\mu^d = \{\mu_j^d\}$, and the buffer size N as its inputs. In this specific example, since $\psi_{ii'}^u = \frac{\mu_j^d}{\mu_i^u}\lambda_{ii'}^d$ and $\psi_{jj'}^d = \frac{\mu_i^u}{\mu_j^d}\lambda_{jj'}^d$, ψ^u and ψ^d are defined by the other inputs.

We first order the states of M_u as $\{1, -1, D_1, D_{-1}, D_Q\}$ and number them from 1 to $I_u = 5$. According to the state transitions for M_u and M_d given in Figure 2, the transition rate matrix of M_u is given as

$$\lambda^u = \begin{bmatrix} -g-p & g & p & 0 & 0 \\ 0 & -p-h & 0 & p & h \\ r & 0 & -r & 0 & 0 \\ 0 & r & 0 & -r & 0 \\ r_Q & 0 & 0 & 0 & -r_Q \end{bmatrix}. \tag{49}$$

The processing rates in states $\{1, -1, D_1, D_{-1}, D_Q\}$ are

$$\mu^u = \begin{bmatrix} \mu_u & \mu_u & 0 & 0 \end{bmatrix}.$$

Similarly, the states of M_d are ordered as $\{1', 0'\}$ and numbered from 1 to $I_d = 2$. The transition rate matrix of M_d is given as

$$\lambda^d = \begin{bmatrix} -p' & p' \\ r & -r' \end{bmatrix}. \tag{50}$$

In states $\{1', 0'\}$ the processing rates of M_d are given as

$$\mu^d = \begin{bmatrix} \mu_d & 0 \end{bmatrix}.$$

5.2 Analysis of the Model

Once these inputs are given, we can specify matrices A_1, A_2, A_3, A_4, A_0, B_0, A_N, B_N and vectors $\mathbf{m}_\Upsilon$, $\mathbf{m}_\Delta$, and $\mathbf{m}_Z$ directly. Once these matrices and vectors are specified, the methodology outlined in the preceding sections yields the desired performance measures directly.

The table given in (51) lists the states, the corresponding processing rates, and the classification of each state in sets Υ, Δ, and Z depending on μ_u and μ_d. In this section only the case $\mu_u > \mu_d$ is discussed in detail.

State M_u	State M_d	α_u	α_d	$\mathbf{m}_S$	S $\mu_1 > \mu_2$	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$
1	$1'$	1	1	$\mu_u - \mu_d$	Υ	Z	Δ
-1	$1'$	2	1	$\mu_u - \mu_d$	Υ	Z	Δ
1	$0'$	1	2	μ_d	Υ	Υ	Υ
-1	$0'$	2	2	μ_u	Υ	Υ	Υ
D_1	$1'$	3	1	μ_d	Δ	Δ	Δ
D_{-1}	$1'$	4	1	μ_d	Δ	Δ	Δ
D_Q	$1'$	5	1	μ_d	Δ	Δ	Δ
D_1	$0'$	3	2	0	Z	Z	Z
D_{-1}	$0'$	4	2	0	Z	Z	Z
D_Q	$0'$	5	2	0	Z	Z	Z

(51)

There are 10 discrete states in the state space. When $\mu_u > \mu_d$, $I_\Upsilon = 4$, $I_\Delta = 3$, and $I_Z = 3$. In this case,

$$\mathbf{m}_\Upsilon = \begin{bmatrix} \mu_u - \mu_d & \mu_u - \mu_d & \mu_u & \mu_u \end{bmatrix},$$

$$\mathbf{m}_\Delta = \begin{bmatrix} \mu_d & \mu_d & \mu_d \end{bmatrix},$$

$$\mathbf{m}_Z = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

For this specific case, the submatrices A_1, A_2, A_3, and A_4 are

$$A_1 = \begin{bmatrix} \frac{-p-g-p'}{\mu_u-\mu_d} & 0 & \frac{r'}{\mu_u-\mu_d} & 0 & \frac{r}{\mu_u-\mu_d} & 0 & \frac{r_Q}{\mu_u-\mu_d} \\ \frac{g}{\mu_u-\mu_d} & \frac{-p-h-p'}{\mu_u-\mu_d} & 0 & \frac{r'}{\mu_u-\mu_d} & 0 & \frac{r}{\mu_u-\mu_d} & 0 \\ \frac{p'}{\mu_u} & 0 & \frac{-p-g-r'}{\mu_u} & 0 & 0 & 0 & 0 \\ 0 & \frac{p'}{\mu_u} & \frac{g}{\mu_u} & \frac{-p-h-r'}{\mu_u} & 0 & 0 & 0 \\ -\frac{p}{\mu_d} & 0 & 0 & 0 & \frac{r+p'}{\mu_d} & 0 & 0 \\ 0 & -\frac{p}{\mu_d} & 0 & 0 & 0 & \frac{r+p'}{\mu_d} & 0 \\ 0 & -\frac{h}{\mu_d} & 0 & 0 & 0 & 0 & \frac{r_Q+r'}{\mu_d} \end{bmatrix}, \tag{52}$$

$$A_2 = \begin{bmatrix} 0 & 0 & \frac{r}{\mu_u} & 0 & -\frac{r'}{\mu_d} & 0 & 0 \\ 0 & 0 & 0 & \frac{r}{\mu_u} & 0 & -\frac{r'}{\mu_d} & 0 \\ 0 & 0 & \frac{r_Q}{\mu_u} & 0 & 0 & 0 & -\frac{r'}{\mu_d} \end{bmatrix}^{\mathrm{T}}, \tag{53}$$

$$A_3 = \begin{bmatrix} 0 & 0 & p & 0 & p' & 0 & 0 \\ 0 & 0 & 0 & p & 0 & p' & 0 \\ 0 & 0 & 0 & h & 0 & 0 & p' \end{bmatrix}, \tag{54}$$

$$A_4 = \begin{bmatrix} -r-r' & 0 & 0 \\ 0 & -r-r' & 0 \\ 0 & 0 & -r_Q-r' \end{bmatrix}. \tag{55}$$

The submatrices for the cases $\mu_u = \mu_d$ and $\mu_u < \mu_d$ can be written similarly.

When $\mu_u \neq \mu_d$, the buffer level does not change when both stages are in down states. Since these states cannot be reached when the buffer is empty or full, $S_0 = \Delta$ and $S_N = \Upsilon$. Therefore $I_{S_0} = I_\Delta = 3$ and $I_{S_N} = I_\Upsilon = 4$.

For the empty buffer process, since M_d is completely starved in all transient states, the matrices A_0 and B_0 for the empty buffer process are

$$A_0 = \begin{bmatrix} -r & 0 & 0 \\ 0 & -r & 0 \\ 0 & 0 & -r_Q - r' \end{bmatrix} \tag{56}$$

and

$$B_0 = \begin{bmatrix} r & 0 & r_Q \\ 0 & r & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{57}$$

Since $S_0 = \Delta$, $E[T_0] = -A_0^{-1}$ and $G_0 = -B_0 A_0^{-1}$.

For the full buffer process, M_u is partially blocked in states $(1, 1')$ and $(-1, 1')$ and completely blocked in states $(1, 1')$ and $(-1, 1')$. Then the matrices A_N and B_N are

$$A_N = \begin{bmatrix} -p\frac{\mu_d}{\mu_u} - g\frac{\mu_d}{\mu_u} - p' & 0 & r' & 0 \\ g\frac{\mu_d}{\mu_u} & -p\frac{\mu_d}{\mu_u} - h\frac{\mu_d}{\mu_u} - p' & 0 & r' \\ p' & 0 & -r' & 0 \\ 0 & p' & g\frac{\mu_d}{\mu_u} & -r' \end{bmatrix} \tag{58}$$

$$B_N = \begin{bmatrix} p\frac{\mu_d}{\mu_u} & 0 & 0 & 0 \\ 0 & p\frac{\mu_d}{\mu_u} & 0 & 0 \\ 0 & h\frac{\mu_d}{\mu_u} & 0 & 0 \end{bmatrix}. \tag{59}$$

Since $S_N = \Upsilon$, $E[T_N] = -A_N^{-1}$, $G_N = -B_N A_N^{-1}$.

5.3 Performance Evaluation

Now since all the input matrices and vectors are determined, the solution methodology outlined in the preceding sections yields the probability densities and the performance measures directly. Namely, inserting these matrices and vectors into Equations (25) and (38) with Equations (44) and (46) yields a system of equations that determine the weight vector $\mathbf{w}$. Then Equations (47) and (48) yield production rate and the expected buffer level.

All the results in this section are validated by simulation. Each model is simulated by using both a continuous flow and also a discrete event simulation model. When the continuous simulation is run for 10^6 events, the percentage error between the analytical production rate and the simulated production rate is less than 10^{-5}. The time required to determine the performance measures by using the general methodology is very short and not affected by the buffer level.

Figures 4 and 5 show that increasing the processing rate of each stage increases the production rate until it reaches its limit. However, the expected buffer level increases with the processing rate of the first stage and it reaches its capacity and decreases with the processing rate of the second stage and it approaches zero. Figure 6 shows that increasing the buffer level increases the production rate and the expected buffer level as expected.

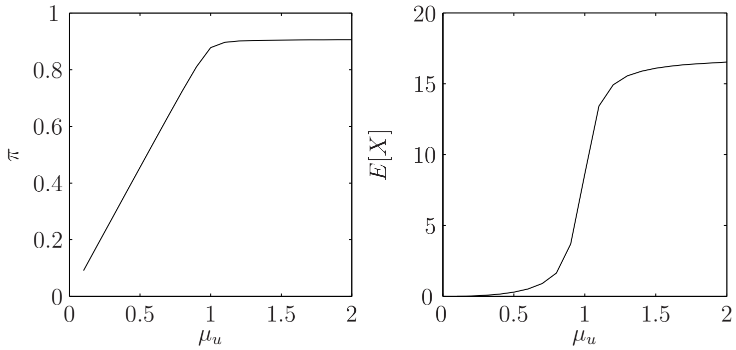

Figure 4: Effect of the processing rate of the upstream station in the model with multiple up and down states ($\mu_d = 1$, $p = 0.005$, $r = 0.15$, $p' = 0.015$, $r' = 0.15$, $g = 0.01$, $h = 0.20$, $r_Q = 0.15$, $N = 17$)

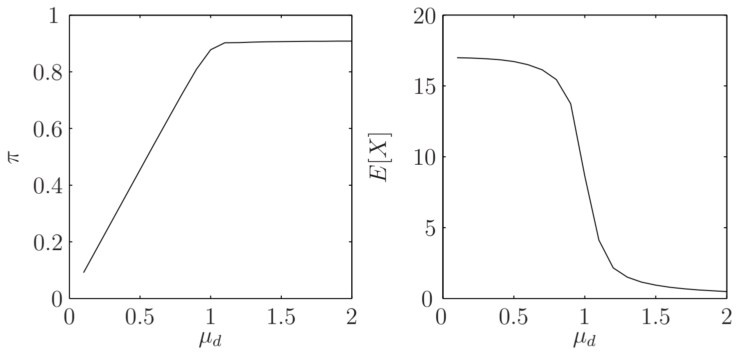

Figure 5: Effect of the processing rate of the downstream station in the model with multiple up and down states ($\mu_u = 1$, $p = 0.005$, $r = 0.15$, $p' = 0.015$, $r' = 0.15$, $g = 0.01$, $h = 0.20$, $r_Q = 0.15$, $N = 17$)

6 Modelling of Various Systems

In this section, we will model four different systems to illustrate the application of our methodology in the analysis of a range of production lines. The first model is a system where each stage has a number of identical machines in parallel. The second model is a system where the up- and down-times of each station are Erlang random variables with different number of stages. Then a model

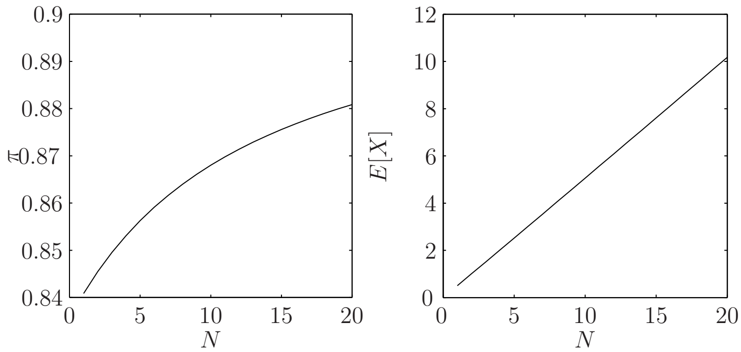

Figure 6: Effect of the buffer capacity in the model with multiple up and down states ($\mu_1 = 1$, $\mu_2 = 1$, $p = 0.005$, $r = 0.15$, $p' = 0.015$, $r' = 0.15$, $g = 0.01$, $h = 0.20$, $r_Q = 0.15$)

of a three station merge system with a shared buffer is discussed. Finally, a model of a system where each stage has a number of machines in series is given. The way these systems are modelled is shown and the inputs are given explicitly for each model.

6.1 A Model with Parallel Machines

We now model a system where M_u has m_u and M_d has m_d identical machines in parallel similar to the one analyzed in Mitra (1988). Each machine is unreliable and has one up and one down state. In the upstream stage, the processing rate of each machine is μ_u and the failure and repair times are exponential random variables with rates p_u and r_u respectively. In the downstream stage, the processing rate of each machine is μ_d and the failure and repair rates are also exponential random variables with rates p_d and r_d respectively.

In this model M_u has $m_u + 1$ and M_d has $m_d + 1$ states. In state i of M_u, i machines are operational and the effective processing rate is $i\mu_u$, $0 \leq i \leq m_u$. Similarly, in state j of M_d, j machines are operational and the effective processing rate is $j\mu_d$, $0 \leq j \leq m_d$.

Accordingly, the possible transitions for M_u are

- from state i to state $i-1$ with rate ip_u for $i = 1, ..., m_u$, and
- from state i to state $i+1$ with rate $(m_u - i)r_u$ for $i = 0, ..., m_u - 1$.

Similarly, possible transitions for M_d are

- from state j to state $j-1$ with rate jp_d for $j = 1, ..., m_d$ and
- from state j to state $j+1$ with rate $(m_d - j)r_d$ for $j = 0, ..., m_d - 1$.

Figure 7 depicts the state transitions for M_u and M_d for a specific case where M_u has $m_u = 3$ machines and M_d has $m_d = 2$ machines in parallel.

The matrices λ^u and λ^d and the vectors μ^u and μ^d for this specific case are given below:

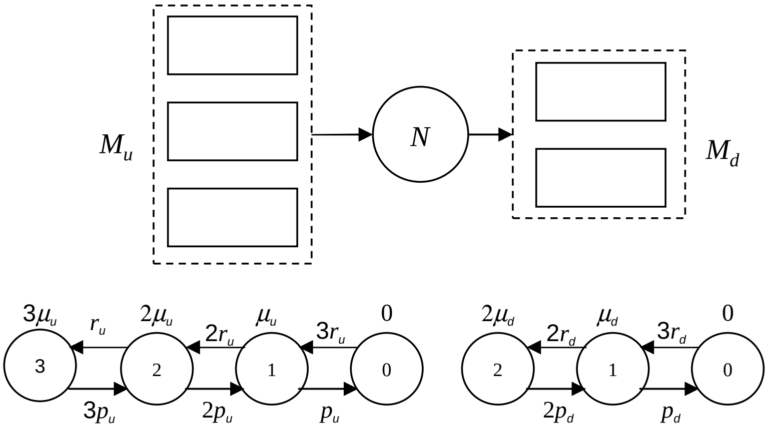

Figure 7: A system with parallel machines

$$\lambda^u = \begin{bmatrix} -3r_u & 3r_u & 0 & 0 \\ p_u & -p_u - 2r_u & 2r_u & 0 \\ 0 & 2p_u & -2p_u - r_u & r_u \\ 0 & 0 & 3p_u & -3p_u \end{bmatrix} \quad (60)$$

where the states are ordered as $\{0, 1, 2, 3\}$. The processing rates in these states are

$$\mu^u = \begin{bmatrix} 0 & \mu_u & 2\mu_u & 3\mu_u \end{bmatrix}.$$

Similarly,

$$\lambda^d = \begin{bmatrix} -2r_d & 2r_d & 0 \\ p_d & -p_d - r_d & r_d \\ 0 & 2p_d & -2p_d \end{bmatrix} \quad (61)$$

where the states are ordered as $\{0, 1, 2\}$. In these states the processing rates of M_d are given as

$$\mu^d = \begin{bmatrix} 0 & \mu_d & 2\mu_d \end{bmatrix}.$$

There are a total of twelve states in the state space. Once these inputs are given, the methodology described above yields the desired performance measures directly. Figure 8 shows the effect of the number of parallel stations on the production rate and the expected buffer level. In this specific case, the production rate of the second stage is kept equal to the production rate of the first stage as the number of parallel stations in the second stage increases. The figures shows that as the number of parallel stations increase both the production rate and the expected buffer level increases.

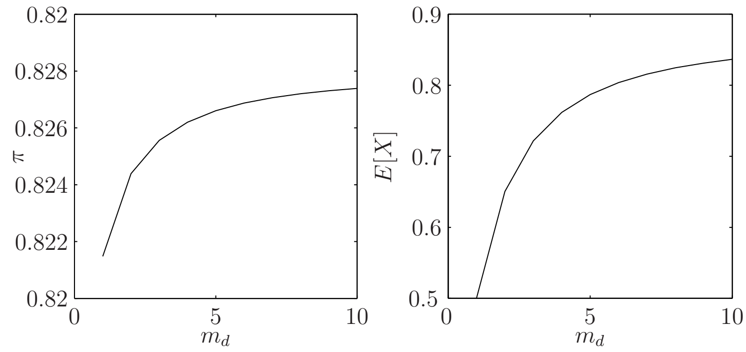

Figure 8: Effect of the number of parallel machines ($\mu_u = 1$, $p_u = 0.01$, $r_u = 0.09$, $m_u = 1$, $\mu_d = \mu_u \frac{m_u}{m_d}$, $p_d = 0.01$, $r_d = 0.09$, $N = 1$)

6.2 A Model with a Shared Buffer

We now consider a three station merge system with a shared buffer. This system was analyzed in detail in (Tan 2001). Helber and Jusic (2004) also analyzes a similar system. In the upstream stage, there are two unreliable machines with processing rates μ_1 and μ_2. In the downstream stage, there is only one machine with processing rate μ_3. The failure and repair rates for each machine are p_i and r_i for $i = 1, 2, 3$. Figure 9 depicts the state transitions for M_u and M_d for this specific case.

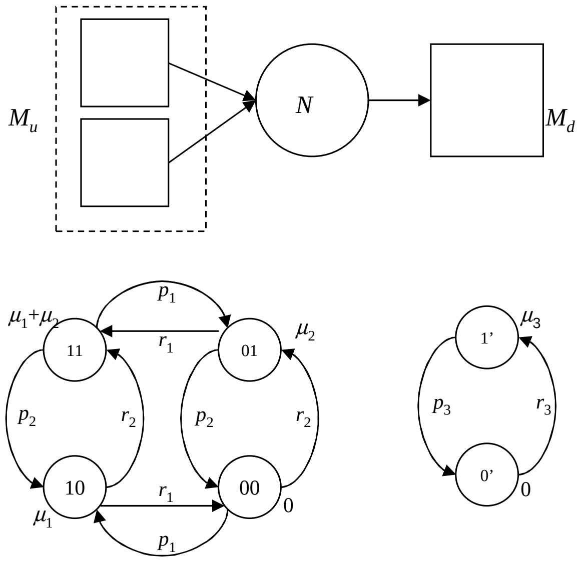

Figure 9: A system with a shared buffer

Similar to the first example, we will specify the matrices λ^u and λ^d and the vectors μ^u and μ^d as the inputs of the solution methodology. The transition rates for M_u are given as

$$\lambda^u = \begin{bmatrix} -p_1 - p_2 & p_2 & p_1 & 0 \\ r_2 & -p_1 - r_2 & 0 & p_1 \\ r_1 & 0 & -p_2 - r_1 & p_2 \\ 0 & r_1 & r_2 & -r_1 - r_2 \end{bmatrix} \tag{62}$$

where the states are ordered as $\{11, 10, 01, 00\}$. The processing rates in these states are

$$\mu^u = \begin{bmatrix} \mu_1 + \mu_2 & \mu_1 & \mu_2 & 0 \end{bmatrix}.$$

Similarly,

$$\lambda^d = \begin{bmatrix} -p_3 & p_3 \\ r_3 & -r_3 \end{bmatrix} \tag{63}$$

where the states are ordered as $\{1, 0\}$. In these states the processing rates of M_d are given as

$$\mu^d = \begin{bmatrix} \mu_3 & 0 \end{bmatrix}.$$

There are eight discrete states in the state space. Once these inputs are given, the methodology described above yields the desired performance measures directly. We compare this case with the results given in (Tan 2001). Since a specific case with hot standby is analyzed in (Tan 2001), the method described above is modified accordingly. Figure 10 shows the effect of μ_3 on the production rate and the expected buffer level obtained by using the methodology given here and the results in (Tan 2001) that are equal to each other.

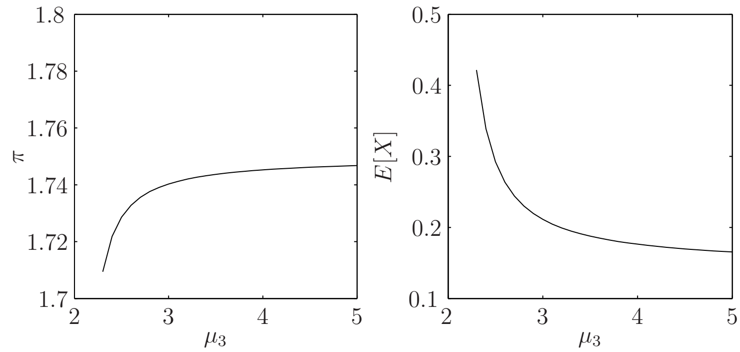

Figure 10: Effect of the processing rate ($\mu_1 = 1.2$, $\mu_2 = 1$, $p_1 = 0.1$, $p_2 = 0.1$, $p_3 = 0.2$, $r_1 = 0.9$, $r_2 = 0.9$, $r_3 = 0.9$, $N = 1$)

6.3 A Model with Erlang Up and Down Times

We now model a production system where the failure and repair times are Erlang-type random variables. We assume that the failure time of M_u is an Erlang random variable with κ_f^u stages.

The expected failure time is $MTTF_u$ and the squared coefficient of variation of the failure time is $scv_f^u = 1/\kappa_f^u$. The repair time of M_u is also an Erlang random variable with κ_r^u stages. The expected failure time is $MTTR_u$ and the squared coefficient of variation of the failure time is $scv_r^u = 1/\kappa_r^u$.

Similarly, failure time of M_d is an Erlang random variable with κ_f^d stages. The expected failure time is $MTTF_d$ and the squared coefficient of variation of the failure time is $scv_f^d = 1/\kappa_f^d$. The repair time of M_d is also an Erlang random variable with κ_r^d stages. The expected failure time is $MTTR_d$ and the squared coefficient of variation of the failure time is $scv_r^d = 1/\kappa_r^d$.

The processing rates of M_u and M_d are μ_u and μ_d respectively. In this model M_u has $\kappa_f^u + \kappa_r^u$ states and M_d has $\kappa_f^d + \kappa_r^d$ states. The states of M_u are indexed from 1 to $\kappa_f^u + \kappa_r^u$ and ordered such that states $i = 1, ..., \kappa_f^u$ are for the up states and states $i = \kappa_f^u + 1, ..., \kappa_f^u + \kappa_r^u$ are for the down states of M_u. Similarly the states of M_d are indexed from 1 to $\kappa_f^d + \kappa_r^d$ and ordered such that states $i = 1, ..., \kappa_f^d$ are for the up states and states $i = \kappa_f^d + 1, ..., \kappa_f^d + \kappa_r^d$ are for the down states of M_u.

The possible transitions for M_u are

- from state i to state $i+1$ with rate $p_u = \kappa_f^u/MTTF_u$, $i = 1, ..., \kappa_f^k$,
- from state i to state $i+1$ with rate $r_u = \kappa_r^u/MTTR_u$, $i = \kappa_f^u + 1, ..., \kappa_f^u + \kappa_r^u - 1$,
- from state $\kappa_f^u + \kappa_r^u$ to state 1 with rate r_u.

Similarly, the possible transitions for M_d are

- from state j to state $j+1$ with rate $p_d = \kappa_f^d/MTTF_d$, $j = 1, ..., \kappa_f^d$,
- from state j to state $i+1$ with rate $r_d = \kappa_r^d/MTTR_d$, $j = \kappa_f^d + 1, ..., \kappa_f^d + \kappa_r^d - 1$,
- from state $\kappa_f^d + \kappa_r^d$ to state 1 with rate r_d.

For example, let us consider a specific case with $\kappa_f^u = 2$, $\kappa_r^u = 2$, $\kappa_f^d = 1$, and $\kappa_r^u = 3$. For this specific system, Figure 11 depicts the state transition diagram.

The matrices λ^u and λ^d and the vectors μ^u and μ^d for this specific case are given below:

$$\lambda^u = \begin{bmatrix} -p_u & p_u & 0 & 0 \\ 0 & -p_u & p_u & 0 \\ 0 & 0 & -r_u & r_u \\ r^u & 0 & 0 & -r_u \end{bmatrix}, \tag{64}$$

$$\mu^u = \begin{bmatrix} \mu_u & \mu_u & 0 & 0 \end{bmatrix},$$

$$\lambda^d = \begin{bmatrix} -p_d & p_d & 0 & 0 \\ 0 & -r_d & r_d & 0 \\ 0 & 0 & -r_d & r_d \\ r^d & 0 & 0 & -r_d \end{bmatrix}, \tag{65}$$

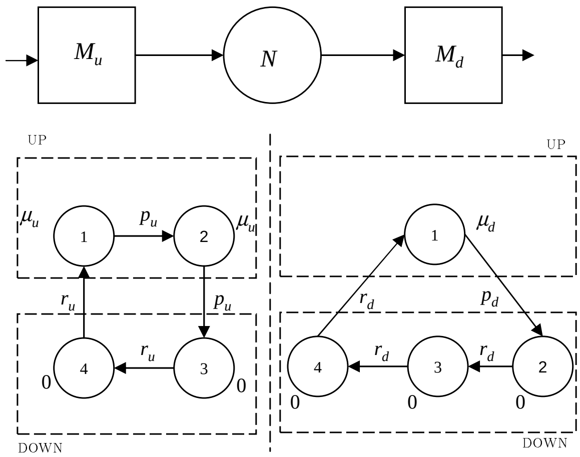

Figure 11: A system with Erlang Up and Down times

$$\mu^d = \left[\begin{array}{cccc} \mu_d & 0 & 0 & 0 \end{array} \right].$$

Figures 12 and 13 show the effects of the failure and repair time variabilities of each stage on the production rate and the expected buffer level. Figure 12 shows that as the coefficient of variation of the failure times of first and the second stages increase, the production rate decreases. On the other hand, a decrease in the variability of the failure time of the upstream machine results in an increase in the expected buffer level. Similarly, Figure 13 shows the effect of the repair time variability of the firs and the second stage on the production rate and the expected buffer level. A decrease in repair time variability of either stage increases the production rate. On the other hand, a decrease of the repair time variability of only the first stage increases the expected buffer level.

6.4 A Model with Series Machines

We now consider a production line where M_u has m_u and M_d has m_d machines in series. The machines are indexed from 1 to $m_u + m_d$. Each machine is unreliable and has one up and one down state. The processing rate of machine k is μ_k. The failure and repair times of machine k are exponential random variables with rates p_k and r_k, $k = 1, ..., m_u + m_d$.

The state of the upstream stage is a vector of length m_u with its ith element is 1 if machine i is operational and 0 otherwise, $1 \leq i \leq m_u$. Similarly, the state of the downstream stage is a vector of length m_d with its jth element is 1 if machine $m_u + j$ is operational and 0 otherwise, $1 \leq j \leq m_d$. Accordingly, M_u has 2^{m_u} states and M_d has 2^{m_d} states.

Since each stage is operational only when all the machines are up, M_u produces at the maximum rate of $\mu_u = \min\{\mu_1, ..., \mu_{m_u}\}$ when all the stations are up and it can not produce if one of the machines is down. Similarly, M_d produces at the maximum rate of $\mu_d = \min\{\mu_{m_u+1}, ..., \mu_{m_u+m_d}\}$ when all the stations are up and it cannot produce when one of the machines is down.

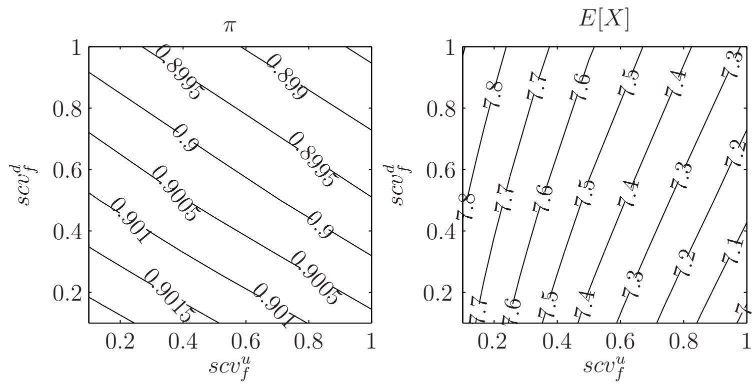

Figure 12: Effect of the failure time variability ($\mu_u = 1$, $\mu_d = 1$, $MTTF_u = 200$, $MTTF_d = 100$, $MTTR_u = 6.67$, $MTTR_d = 10$, $N = 10$)

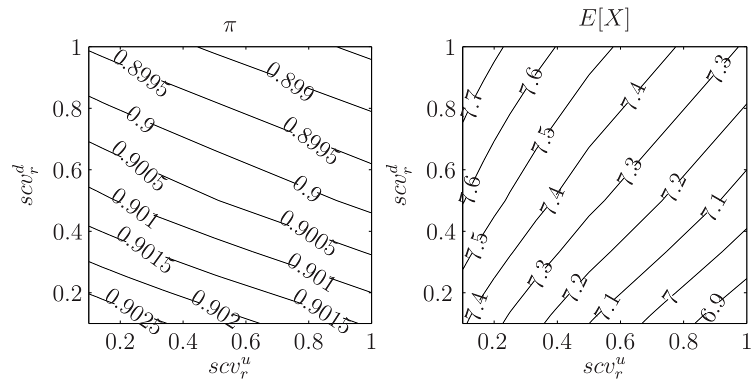

Figure 13: Effect of the repair time variability ($\mu_u = 1$, $\mu_d = 1$, $MTTF_u = 200$, $MTTF_d = 100$, $MTTR_u = 6.67$, $MTTR_d = 10$, $N = 10$)

When all the machines of M_u are up, each machine can fail with rate $p_i \frac{\mu_u}{\mu_i}$, $i = 1, ..., m_u$ due to operational failures. Similarly, when one machine is down, none of the other up machines can fail since they will be forced to stop due to the down machine. As a result, the only possible transition when machine k is down is the repair of machine k with rate r_k. Therefore although there are 2^{m_u} states for M_u, only $m_u + 1$ of them will be non-transient. The case for M_d is similar.

Figure 14 depicts the state transitions for M_u and M_d for a specific case where M_u has 3 machines and M_d has 2 machines in series.

The matrices λ^u and λ^d and the vectors μ^u and μ^d for this specific case are given below:

$$\lambda^u = \begin{bmatrix} -\mu_u(\frac{p_1}{\mu_1} + \frac{p_2}{\mu_2} + \frac{p_3}{\mu_3}) & p_1\frac{\mu_u}{\mu_1} & p_2\frac{\mu_u}{\mu_2} & p_3\frac{\mu_u}{\mu_3} \\ r_1 & -r_1 & 0 & 0 \\ r_2 & 0 & -r_2 & 0 \\ r_3 & 0 & 0 & -r_3 \end{bmatrix} \tag{66}$$

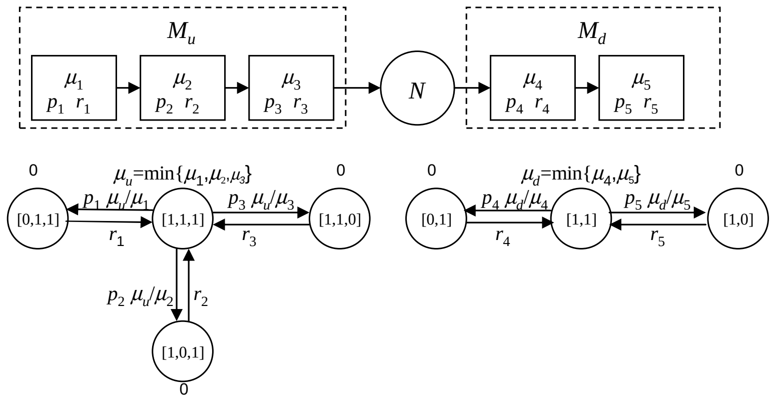

Figure 14: Modelling of a system with series machines for analysis by using the general methodology

where the states are ordered as $\{(1,1,1),(0,1,1),(1,0,1),(1,1,0)\}$. The processing rates in these states are

$$\mu^u = \left[\begin{array}{cccc} \mu_u & 0 & 0 & 0 \end{array} \right]$$

where $\mu_u = \min\{\mu_1, \mu_2, \mu_3\}$. Similarly

$$\lambda^d = \left[\begin{array}{ccc} -\mu_d(\frac{p_4}{\mu_4} + \frac{p_5}{\mu_5}) & p_4\frac{\mu_d}{\mu_4} & p_5\frac{\mu_d}{\mu_5} \\ r_4 & -r_4 & 0 \\ r_5 & 0 & -r_5 \end{array} \right] \quad (67)$$

where the states are ordered as $\{(1,1),(1,0),(0,1)\}$. In these states the processing rates of M_d are given as

$$\mu^d = \left[\begin{array}{ccc} \mu_d & 0 & 0 \end{array} \right]$$

where $\mu_d = \min\{\mu_4, \mu_5\}$. There are a total of twelve states in the state space. Once these inputs are given, the methodology described above yields the desired performance measures directly.

Consider the problem of locating a finite buffer in a continuous material flow production line with no interstation buffers. Once the buffer is located between machine k and $k+1$, the line is divided into two stages. The resulting two-stage system can be analyzed by using the methodology outlined above. Figure 15 shows the effect of the buffer placement on the production rate for a production line with ten identical stations. As expected, for this homogeneous system placing the buffer in the middle, between Machine 5 and 6 maximizes the production rate.

However, when the machines are not identical, the buffer location that maximizes the production rate can be different. Figure 16 shows the effect of the buffer placement on the production rate for a production line with ten non-identical stations. In this case, placing the buffer between Machine 5 and 6 maximizes the production rate.

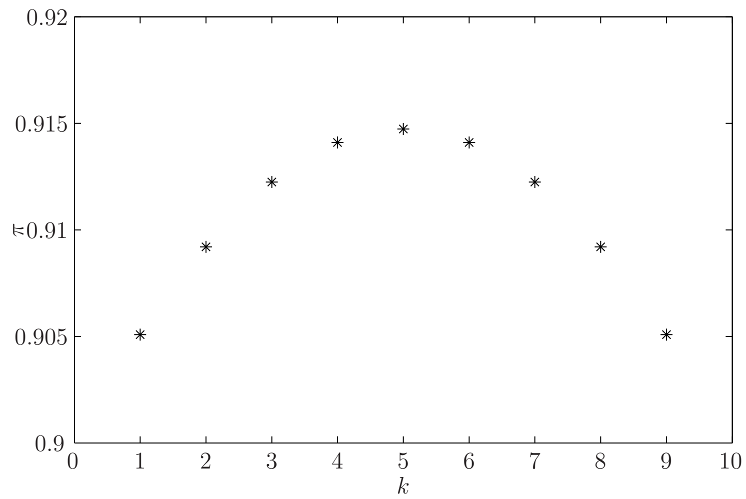

Figure 15: Effect of the buffer placement on the production rate ($\mu_i = 1$, $p_i = 0.01$, $r_i = 0.9$, $i = 1, ..., 10$, $N = 1$)

7 Conclusion

We presented a general methodology to analyze continuous-flow material flow two stage-single buffer production systems. The method handles general Markovian transitions and different processing rates associated with each state for both stages. The run time of the method is very fast and not affected by the buffer size.

A wide range of models can be analyzed by our methodology directly by determining the transition rates of each stage and the flow rates associated with the discrete states of each stage. We used the methodology presented in this study to model and analyze various single buffer continuous flow systems including the systems where each stage has a number of identical machines in parallel or in series, systems where the up- and down-times of each station are Erlang random variables with different number of stages, and a model of a three-station merge system with a shared buffer.

In addition to the production systems, our methodology can also be used in performance evaluation of computer and telecommunication systems. Since the operation-dependent failure mechanism differentiates the models of production and computer/telecommunication models, setting the operation dependent failure rates equal to the original rates in our methodology allows us to use the same tool in the performance evaluation of computer and telecommunication systems.

Therefore we propose our model as a general tool to model and analyze single buffer fluid flow systems.

Acknowledgement

The first author acknowledges the support from TÜBA and TÜBİTAK. The second author acknowledges the support from the Singapore-MIT Alliance and the General Motors Research and

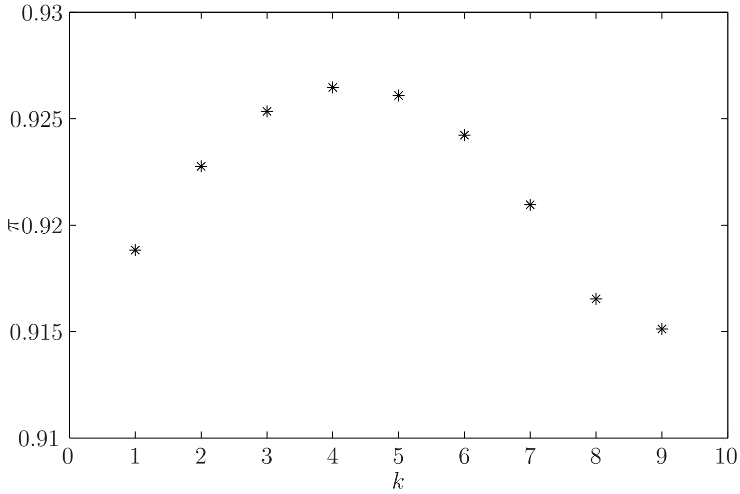

Figure 16: Effect of the buffer placement on the production rate ($p_i = 0.01$, $r_i = 0.9$, $i = 1 : 10$, $\mu_j = 1$, $j = 1 : 8$, $\mu_k = 4$, $k = 9, 10$, $N = 1$)

Development Center.

References

Ahn, S., J. Jeon, and V. Ramaswami (2005). Steady state analysis for finite fluid flow models using finite QBDs. *Queueing Systems 49*, 223259.

Ahn, S. and V. Ramaswami (2003). Fluid flow models and queues: a connection by stochastic coupling. *Stochastic Models 19*(3), 325–348.

Altıok, T. (1985). On the phase-type approximation of general distributions. *AIIE Transactions 17*, 110–116.

Anick, D., D. Mitra, and M. Sondhi (1982). Stochastic theory of a data handling system with multiple sources. *Bell System Technology Journal 61*, 1871–1894.

Bertsimas, D. (1990). An analytical approach to a general class of G/G/s queueing systems. *Operations Research 38*(1), 139–155.

Blake, I. and W. Lindsey (1973). Level-crossing problems for random processes. *IEEE Transactions on Information Theory 19*(3), 295–315.

Brill, P. H. (1978). An embedded level crossing technique for dams and queues. *Journal of Applied Probability 16*(1), 174–186.

Dallery, Y. (1994). On modeling failure and repair times in stochastic models of manufacturing systems using generalized exponential distributions. *Queueing Systems Theory Applications 15*(1-4), 199–209.

Diamantidis, A., C. Papadopoulos, and M. Vidalis (2004). Exact analysis of a discrete material three-station one-buffer merge system with unreliable machines. *International Journal of Production Research 42*(4), 651–675.

Dubois, D. and J. P. Forestier (1982). Productivité et en-cours moyens d'un ensemble de deux machines séparées par une zone de stockage. *RAIRO Automatique 16*, 105–132.

Elwalid, A. and D. Mitra (1991). Analysis and design of rate-based congestion control of high speed networks, i: stochastic fluid models, access regulation. *Queueing Systems Theory Applications 9*, 19–64.

Gershwin, S. B. and S. Fallah-Fini (2007). A general model and analysis of a discrete two-machine production line. In *Analysis of Manufacturing Systems 2007*.

Gershwin, S. B. and I. C. Schick (1980). Continuous model of an unreliable two-machine material flow system with a finite interstage buffer. Report LIDS-R-1039, MIT Laboratory for Information and Decision Systems.

Helber, S. and H. Jusic (2004). A new decomposition approach for non-cyclic continuous material flow lines with a merging flow of material. *Annals of Operations Research 125*, 117139.

Levantesi, R., A. Matta, and T. Tolio (2003). Performance evaluation of continuous production lines with machines having different processing times and multiple failure modes. *Performance Evaluation 51*, 247–268.

Mitra, D. (1988). Stochastic theory of a fluid model of multiple failure-susceptible producers and consumers coupled by a buffer. *Advances in Applied Probability 20*, 646–676.

Özdoğru, U. and T. Altıok (2003). Analysis of two-valve fluid flow systems with general repair times. In *Analysis and Modeling of Manufacturing Systems*, pp. 255–288. Kluwer Publications.

Patchong, A. and D. Willaeys (2001). Modeling and analysis of an unreliable flow line composed of parallel-machine stages. *IIE Transactions 33*, 559–568.

Poffe, A. and S. Gershwin (2005). Integrating quality and quantity modelling in a production line. Technical Report ORC-377-05, Massachusetts Institute of Technology Operations Research Center Working Paper Series.

Serucola, B. (2001). A finite buffer fluid queue driven by a markovian queue. *Queueing Systems 38*, 213–220.

Soares, A. D. S. and G. Latouche (2006). Matrix-analytic methods for fluid queues with finite buffers. *Performance Evaluation 63*, 295314.

Tan, B. (2001). A three station continuous materials flow merge system with unreliable stations and a shared buffer. *Mathematical and Computer Modelling 33*(8-9), 1011–1026.

Tan, B. and S. B. Gershwin (2007). Modelling and analysis of markovian continuous flow production systems with a finite buffer: A general methodology and applications. Technical Report ORC-381-07, Massachusetts Institute of Technology Operations Research Center Working Paper Series.

Tempelmeier, H. and M. Burger (2001). Performance evaluation of unbalanced flow lines with general distributed processing times, failures and imperfect production. *IIE Transactions 33*, 293 – 302.

Wijngaard, J. (1979). The effect of interstage buffer storage on the output of two unreliable production units in series with different production rates. *AIIE Transactions 11*(1), 42–47.

Yeralan, S., W. Franck, and M. A. Quasem (1986). A continuous materials flow production line model with station breakdown. *European Journal of Operational Research 27*, 289–300.

Yeralan, S. and B. Tan (1997). A station model for continuous materials flow production. *International Journal of Production Research 35*(9), 2525–2541.