*Control Eng. Practice*, Vol. 1, No. 1, pp. 147-152, 1993
Printed in Great Britain. All rights reserved.

0967-0661/93 $6.00 + 0.00


# ON ASYMPTOTICALLY RELIABLE CLOSED SERIAL PRODUCTION LINES

J.-T. Lim*,[1] and S.M. Meerkov*,[2]

**Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Taejon, Korea*
***Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109-2122, USA*

**Abstract.** A problem of analysis and design of serial production lines, closed with respect to the number of carriers available for parts transportation between operations, is formulated. An asymptotic solution is given for two machines-two buffers systems. A case study of a paint shop operation at an automobile assembly plant is described. It is shown that optimization of the system with respect to the number of carriers and the capacity of the feedback buffer may lead to a substantial improvement of system's performance.

**Key Words.** Production systems; Markov processes; asymptotic analysis, production rate and work-in-process calculations.

## 1. INTRODUCTION

Consider a manufacturing system defined by the following assumptions:

(i) The system consists of $M$ machines $m_i, i = 1, \ldots, M$, arranged in the consecutive order, and $M-1$ buffers, $B_i$, separating each two machines, $m_i$ and $m_{i+1}$.

(ii) The machines have identical cycle time $T$. The time axis is slotted with the slot duration $T$. Each machine begins its operation at the beginning of the time slot.

(iii) Each buffer is characterized by its capacity, $N_i$, $i = 1, \ldots M-1$, where $N_i$ is a positive integer.

(iv) Machine $m_i$ is starved during a time slot if buffer $B_{i-1}$ is empty at the beginning of this time slot; machine $m_i$ is blocked during a time slot if at the beginning of this time slot buffer $B_i$ is full and machine $m_{i+1}$ is either down or blocked. Machine $m_1$ is never starved, machine $m_M$ is never blocked.

(v) Machine $m_i$, being not blocked and not starved, produces a part during any time slot with probability $q_i = 1 - \epsilon k_i$ and fails to do so with probability $\epsilon k_i, i = 1, \ldots, M$, where $0 < \epsilon << 1$ and $k_i > 0$ is independent of $\epsilon$. The $k_i$'s are called the loss parameters.

Manufacturing systems defined by assumptions (i)-(v) are referred to as asymptotically reliable, *open* serial production lines. An asymptotic method for their analysis and design has been developed by Lim *et al.* (1989, 1990). In many practical situations, however, serial production lines are *closed*, i.e., have a feedback loop with respect to carriers on which the parts (jobs) are transported from one machine to another. This is, in particular, the case in assembly and painting operations in the automobile industry where car bodies and engine blocks are transferred between operations on carriers, and the number of carriers in the system is constant. To account for this situation, introduce the following assumption:

(vi) The jobs are transported within the system on carriers. Each job is placed on a carrier at the input of machine $m_1$ and is removed from the carrier at the output of machine $m_M$. Empty carriers are returned to the empty carrier buffer, $B_M$, and are supplied to the input of $m_1$ instantaneously, given that $B_M$ is not empty. The capacity of $B_M$ is $N_M$. The total

[1]Supported by the Korea Science and Engineering Foundation under grant No. 880115.
[2]Supported by the NSF grant No. DDM-9105086.

number of carriers in the system is $S$, where $M \leq S \leq \sum_{i=1}^{M} N_i$.

Manufacturing systems defined by (i)-(vi) are referred to as asymptotically realiable *closed* serial production lines. The block diagram of such a system is shown in Fig. 1.

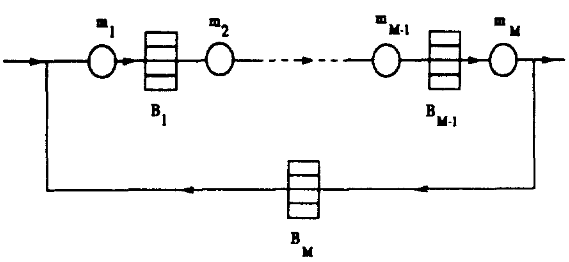


Fig. 1. Closed serial production line.

Performance of closed lines can be characterized by their production rate, $PR_c$, i.e., the average number of jobs produced in the steady state by the last machine, $m_M$, and the work-in-process, $WIP_c$, i.e., the average number of jobs in the system at the steady state. The problem of analysis of these lines is formulated as follows: Given $k_1, N_1, \ldots, k_M, N_M$, and $S$, find $PR_c(k_1, N_1, \ldots, k_M, N_M, S)$ and $WIP_c(k_1, N_1, \ldots k_M, N_M, S)$. The problem of design is: Given $k_1, N_1, \ldots, k_{M-1}, N_{M-1}, k_M$ find the smallest $S$ and $N_M$ so that

$$\begin{aligned} &PR_c(k_1, N_1, \ldots, k_M, N_M, S) \\ &= PR_0(k_1, N_1, \ldots, k_{M-1}, N_{M-1}, k_M)\ , \\ &WIP_c(k_1, N_1, \ldots, k_M, N_M, S) \\ &= WIP_0(k_1, N_1, \ldots, k_{M-1}, N_{M-1}, k_M)\ , \end{aligned} \tag{1.1}$$

i.e., find the condition under which the feedback does not impair the performance of the line ($PR_0$ and $WIP_0$ in (1.1) denote the production rate and the work-in-process of the open line, respectively).

The purpose of this paper is to give a solution to these problems for $M = 2$. Specifically, it is shown below that the problem of analysis of closed lines can be reduced to that of open ones (Section 3). Thus, the methods of Lim *et al.* (1990) become applicable to closed serial production systems. For the purpose of design, it is shown how to choose $S$ and $N_2$ so that (1.1) is asymptotically satisfied (Section 4). Finally, the results obtained are applied to a case study of a paint shop operation at a modern automobile assembly plant, and it is shown that a substantial improvement can be achieved by the optimal choice of $S$ and $N_2$ (Section 5).

The problem at hand belongs to the area of closed Markovian queueing systems, which has been studied in many publications. Most of them, however, address non-blocking systems with infinite buffers (see, for instance, Gordon and Newell (1967a), Posner and Bernholtz (1968), Kobayashi (1976a,b), Reiser and Lavenberg (1980), Boxma *et al.* (1984), McKenna and Mitra (1984), Suri and Diehl (1986)). For finite buffers, the joint steady state distribution of buffer occupancy has been calculated by Gordon and Newell (1976b). In Akyildiz (1987a,b), an approximate analysis has been carried out using a state space reduction technique. An application of the mean value analysis has been reported in Akyildiz (1988). Papers of Onvural and Perros (1989a,b) present another approximation technique based on a curve fitting approach.

In spite of these achievements, analytic methods for performance evaluation of closed Markovian queueing systems are still missing. This paper is intended to contribute to this end.

## 2. PRELIMINARIES

Consider an open serial production line defined by (i)-(v) with $M = 2$ and the buffer capacity $N$. Let $h(n)$, $n = 0, 1, \ldots$, be the occupancy of the buffer at the beginning of slot $[n, n+1)$. Introduce

$$v_j(n) = Prob\{h(n) \geq j\}, \quad j = 1, \ldots, N\ . \tag{2.1}$$

Let $v_j$ be the steady state value of $v_j(n)$, i.e., the probability that there are at least $j$ parts in the buffer at the steady state. Define the following function

$$Q(\alpha, N) = \frac{1-\alpha}{1-\alpha^N}, \alpha \in R_+\ . \tag{2.2}$$

**Theorem 2.1** (Lim *et al.* 1990): The performance of open serial production lines defined by (i)-(v) with $M = 2$ is as follows:

(a) The steady state distribution of buffer occupancy is:

$$\begin{aligned} v_1 &= 1 - \epsilon k_1 Q\left(\frac{k_2}{k_1}, N\right) + O(\epsilon^2) \\ v_N &= Q\left(\frac{k_1}{k_2}, N\right) v_1 + O(\epsilon) \\ v_i &= Q^{-1}\left(\frac{k_1}{k_2}, N - i + 1\right) v_N + O(\epsilon), \\ & \quad i = 2, \ldots N-1\ . \end{aligned} \tag{2.3}$$

(b) The production rate is

$$\begin{aligned} &PR_0(k_1, N, k_2) \\ &= 1 - [k_2 + k_1 Q(\frac{k_2}{k_1}, N)]\epsilon + O(\epsilon^2) \\ &= 1 - [k_1 + k_2 Q(\frac{k_1}{k_2}, N)]\epsilon + O(\epsilon^2) \end{aligned} \tag{2.4}$$

Formula (2.3) can be used to calculate the system's work-in-process. Indeed, since

$$WIP_0(k_1, N, k_2) = \sum_{i=1}^{N} i(v_i - v_{i+1}) = \sum_{i=1}^{N} v_i ,$$

from (2.3), it follows that

$$WIP_0(k_1, N, k_2) = \frac{NQ(\frac{k_1}{k_2}, N)k_2 - k_1}{k_2 - k_1} + O(\epsilon) . \quad (2.5)$$

It is shown below that the performance evaluation of closed lines can be reduced to that of open lines, where the effective buffer capacity, $N_e$, depends on the relationship between $N_1, N_2$ and $S$.

## 3. ANALYSIS

**Theorem 3.1:** The performance of closed serial production lines defined by (i)-(vi) with $M = 2$ is characterized as follows:

$$PR_c(k_1, N_1, k_2, N_2, S) = PR_0(k_1, N_e, k_2) + O(\epsilon^2) , \quad (3.1)$$

$$WIP_c(k_1, N_1, k_2, N_2) = \begin{cases} \max(0, S - N_2 - 1) \\ +WIP_0(k_1, N_e, k_2) + O(\epsilon), \\ \qquad \text{if } N_1 \le N_2 , \\ S - [\max(0, S - N_1 - 1) \\ +WIP_0(k_2, N_e, k_1)] + O(\epsilon), \\ \qquad \text{if } N_1 > N_2 , \end{cases} \quad (3.2)$$

where $N_e$, the effective buffer size, is given by

$$N_e = \begin{cases} S - 1, & \text{for } 2 \le S \le \min(N_1, N_2), \\ \min(N_1, N_2), & \text{for } \min(N_1, N_2) < S \le \max(N_1, N_2), \\ N_1 + N_2 - S + 1, & \text{for } \max(N_1, N_2) < S \le N_1 + N_2. \end{cases} \quad (3.3)$$

**Proof:** See the Appendix.

Theorem 3.1 can be interpreted as follows:

(i). Since $N_e \le N_1$ and $PR_0$ and $WIP_0$ are monotonically increasing functions of $N_e$, the performance of a closed line cannot supersede that of the corresponding open line. If $S$ and $N_2$ are chosen so that $N_e < N_1$, the feedback impairs the performance of the open line.

(ii). Like the open lines, the closed lines under consideration are equivalent to a single, aggregated machine characterized, in isolation, by the parameter

$$q_{aggregation} = 1 - [k_2 + k_1 Q(\frac{k_2}{k_1}, N_e)]\epsilon + O(\epsilon^2) = 1 - [k_1 + k_2 Q(\frac{k_1}{k_2}, N_e)]\epsilon + O(\epsilon^2) ,$$

(iii). When $S = 2$ and $S = N_1 + N_2$, the aggregated loss parameter is the sum of the losses of both machines no matter how large $N_1$ is:

$$k_{aggregation} = k_1 + k_2 .$$

For $2 < S < N_1 + N_2$, function $Q(\alpha, N_e)$ describes the attenuation of perturbations (failures) introduced by machines.

(iv). From (3.1),

$$PR_c(k_1, N_1, k_2, N_2, S) = PR_c(k_1, N_2, k_2, N_1, S) .$$

However, from (3.2),

$$WIP_c(k_1, N_1, k_2, N_2, S) \le WIP_c(k_1, N_2, k_2, N_1, S) \text{ , if } N_1 \le N_2 .$$

This means that large $S$ and $N_1$ with small $N_2$ will result in large $WIP$ whereas large $S$ and $N_2$ with small $N_1$ will result in small $WIP$ but the same production rate.

## 4. DESIGN

**Proposition 4.1:** Under assumption (i)-(v) with $M = 2$, the asymptotic solution of the design problem (1.1), i.e., the smallest $S$ and $N_2$ which, for given $k_1, N_1$ and $k_2$, guarantee that

$$\begin{aligned} &PR_c(k_1, N_1, k_2, N_2, S) \\ &= PR_0(k_1, N_1, k_2)g + O(\epsilon^2) , \\ &WIP_c(k_1, N_1, k_2, N_2, S) \\ &= WIP_0(k_1, N_1, k_2)g + O(\epsilon^2) , \end{aligned} \quad (4.1)$$

is given by

$$S = N_2 = N_1 + 1 . \quad (4.2)$$

**Proof:** Follows directly from Theorem 3.1.

From (3.1)-(3.3), it is clear that equalities (4.1) also take place for

$$N_2 > N_1 + 1, \; N_1 + 1 < S \le N_2 . \quad (4.3)$$

This solution, however, is inferior to (4.2) since (4.3) results in the same performance but requires a larger number of carriers and feedback capacity.

## 5. CASE STUDY

Lim *et al.* (1990) analyzed a paint shop operation at an automobile assembly plant as an open serial production line, although in reality the system is closed with respect to the carriers (skids) (see Fig. 4 of Lim *et al.* (1990)). In Lim *et al.* (1989), an improvement measure for the paint shop has been developed, again based on the open loop approach. In these publications, the closed loop effects have been taken into account by heuristic considerations. Below, we apply the results of Sections 3 and 4 and develop an improvement measure based on the closed lines approach.

As it has been shown in Lim *et al.* (1990), the paint shop description can be reduced to a closed serial production line shown in Fig. 2. In this figure, P.O. and F.O. stand for Preparation and Final Operations, respectively. The rectangles are the accumulators. Max and min numbers in the accumulators give the maximal capacity of the accumulator and its minimal occupancy necessary to sustain the planned production rate of 63 jobs/hour. The difference between max and min numbers gives the capacity of buffers, i.e., $N_1 = 26$ and $N_2 = 76$. The numbers above each element in Fig. 2 represent the numbers of skids in the nominal conditions. The total number of skids in the system is 650. Since the number of skids in the accumulators are just enough to sustain 63 jobs/hour, the effective $S_e$ (i.e., the total number of skids in two buffers whose capacities are 26 and 76) turns out to be 2.

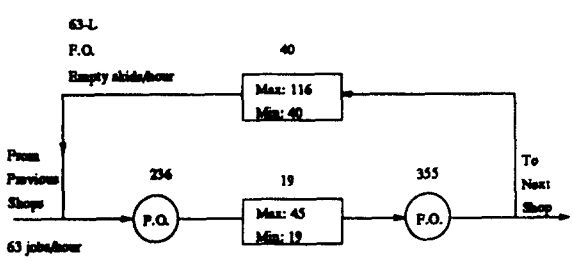


Fig. 2. Simplified paintshop block diagram

Thus, using Theorem 3.1,

$$\begin{aligned} PR_c &= 1 - (k_1 + k_2)\epsilon + O(\epsilon^2) \text{ (jobs/cycle)} \\ \text{or } PR_c &= 63 - (L_{P.O.} + L_{F.O.}) \text{ (jobs/hour)} \end{aligned} \tag{5.1}$$

where $L_{P.O.}$ and $L_{F.O.}$ are the average losses in P.O. and F.O., respectively, and $\epsilon k_1$ and $\epsilon k_2$ are defined as

$$\epsilon k_1 = \frac{L_{P.O.}}{63}, \epsilon k_2 = \frac{L_{F.O.}}{63} .$$

The values of $L_{P.O.}$ and $L_{F.O.}$ for five consecutive monthly periods are given in Table 1. Using these values, production rate estimate (5.1) has been calculated and compared with the corresponding periods (Table 2). As it follows from this table, the accuracy of the prediction is sufficiently high, with the exception of period 2. It turned out that during this period a new car model has been introduced in the production, and it is supposed that additional perturbations, not included in the model, played a crucial role.

Turning now to the improvement measures, from Proposition 4.1 it is concluded that if $S_e$ is chosen optimally, i.e.,

$$S_e = N_1 + 1 = 27 ,$$

the production rate is as shown in Table 3. Thus, if the total number of skids in the system is

$$S = S_e + 650 = 27 + 650 = 677 ,$$

the production rate is increased by 10.5%.

Combined with the improvement measures suggested in Lim *et al.* (1989), where unbalanced times for operations in P.O. and F.O. have been assigned, the suggested above choice of $S_e$ can bring the production rate estimate up to the planned value of 63 jobs/hour (17.2% improvement).

## 6. CONCLUSIONS

(i) Analysis of asymptotically reliable closed serial production lines with 2 machines and 2 buffers can be reduced to that of the corresponding open lines.

(ii) The effective buffer capacity under this reduction is defined by the number of carriers available in the system and the capacity of the buffers in both feedforward and feedback paths.

(iii) The optimal choice of the number of carriers in the system and the capacity of the feedback buffer can lead to a substantial improvement in the system's performance.

## 7. REFERENCES

Akyildiz, I.F. (1987a). Exact Product Form Solution for Queueing Networks with Blocking, *IEEE Trans. on Computer*, **36**, 122-125.

Akyildiz, I.F. (1987b). On the Exact and Approximate Throughput Analysis of Closed Queueing Networks with Blocking, *IEEE Trans. on Software Eng.*, **14**, 62-70.

Akyildiz, I.F. (1988). Mean Value Analysis for Blocking Queueing Networks, *IEEE Trans. on Software Eng.*, **14**, 418–428.

Boxma, O.J., F.P. Kelly, and A.G. Konhein. (1984). The Product From for Sojourn Time Distributions in Cyclic Exponential Queues, *J. Assoc. Comp. Mach.*, **31**, 128–133.

Table 1. Average losses (jobs/hr)

| Periods | Time Period 1 | Time Period 2 | Time Period 3 | Time Period 4 | Time Period 5 |
|---|---|---|---|---|---|
| P.O. | 3.77 | 3.94 | 4.46 | 3.25 | 2.95 |
| F.O. | 6.18 | 7.38 | 7.01 | 6.59 | 6.14 |

Table 2. Actual and Estimated Production Rates (jobs/hr)

| Periods | Time Period 1 | Time Period 2 | Time Period 3 | Time Period 4 | Time Period 5 |
|---|---|---|---|---|---|
| Acutal PR | 53.50 | 43.81 | 51.27 | 54.28 | 55.89 |
| Estimated PR | 53.05 | 51.68 | 51.53 | 53.16 | 53.91 |
| Error (%) | 0.8 | 18.0 | 0.5 | 2.1 | 3.5 |

Table 3. Expected Production Rates (jobs/hr)

| Periods | Time Period 1 | Time Period 2 | Time Period 3 | Time Period 4 | Time Period 5 |
|---|---|---|---|---|---|
| Expected PR | 59.23 | 59.06 | 58.54 | 59.76 | 60.05 |

Gordon, W.J., and F.G. Newell. (1967a). Closed Queueing Sytems with Exponential Servers, *Opns. Res.*, **15**, 254–265.

Gordon, W.J., and F.G. Newell. (1967b). Closed Queueing Sytems with Restricted Queues, *Opns. Res.*, **15**, 266–277.

Kobayashi, H. (1974a). Application of the Diffusion Approximation to Queueing Networks I: Equilibrium Distributions, *J. Assoc. Comp. Mach.*, **21**, 316–328.

Kobayashi, H. (1974b). Application of the Diffusion Approximation to Queueing Networks II: Nonequilibrium Distributions and Applications to Computer Modeling, *J. Assoc. Comp. Mach.*, **21**, 459–469.

Lim, J.-T., S.M. Meerkov, and F. Top. (1989). Analysis and Synthesis of Asynchoronous, Asymptotically Realiable, Asymptotically Controllable Serial Production Lines, *Proc. 28th IEEE CDC*, **1**, 60–64, Tampa, FL.

Lim, J.-T., S.M. Meerkov, and F. Top. (1990). Homogeneous, Asymptotically Realiable Serial Production Lines: Theory and Case Study, *IEEE Trans. Automat. Contr.*, **AC-35**, 524–534.

Mckenna, J. and D. Mitra. (1984). Asymptotic Expansions and Integral Representations of Moments of Queue Lengths in Closed Markovian Networks, *J. Assoc. Comp. Mach.*, **31**, 346–360.

Onvural, R.O. and H.G. Perros. (1989a). Some Equivalencies Between Closed Queueing Networks with Blocking, *Performance Evaluation*, **9**, 111–118.

Onvural, R.O. and H.G. Perros. (1989b). Approximate Throughout Analysis of Cyclic Queueing Networks with Finite Buffers, *IEEE Trans. on Software Eng.*, **15**, 800–808.

Posner, M. and B. Bernholtz. (1968). Closed Finite Queueing Networks with Time Lags and with Several Classes of Units, *Opns. Res.*, **16**, 977–985.

Reiser M. and S.S. Lavenberg. (1980). Mean-Value Analysis of Closed Multichain Queueing Networks, *J. Assoc. Comp. Mach.*, **27**, 313–322.

Suri, R. and G.W. Diehl. (1986). A Variable Buffer-Size Model and its Use in Analyzing Closed Queueing Networks with Blocking, *Management Science*, **32**, 206–224.

## A. APPENDIX

**Proof of Theorem 3.1:** If $2 \le S \le \min(N_1, N_2)$, the occupancy of buffer $B_1$ is at most $S$. Thus, the effective capacity of $B_1$ is reduced to $S$. Therefore, states $v_i(n), i = 1, \ldots, S$, defined in (2.1) can be described by the following equations:

$$\begin{aligned} v_1(n+1) &= v_1(n) + (1 - v_i(n))q_1 + [v_2(n) \\ &\quad + (v_1(n) - v_2(n))q_1]\, q_2 - v_1(n)q_2 \\ v_i(n+1) &= v_i(n) + (v_{i-1}(n) - v_i(n))q_1 \\ &\quad + [v_{i+1}(n) + (v_i(n) - v_{i+1}(n))q_1]\, q_2 \\ &\quad - [v_i(n) + (v_{i-1}(n) - v_i(n))q_1]\, q_2\, , \\ v_S(n+1) &= v_S(n) + (v_{S-1}(n) - v_S(n))q_1 \\ &\quad - [v_S(n) + (v_{S-1}(n) - v_S(n))q_1]\, q_2\, . \end{aligned} \tag{A.1}$$

In the steady state, we obtain

$$\begin{aligned} (1 - v_S)q_1 - v_1 q_2 &= 0 \\ (1 - v_{S-1})q_1 + [v_S + (v_{S-1} - v_S)q_1]\, q_2 - v_1 q_2 &= 0 \end{aligned} \tag{A.2}$$

and

$$v_{S-1} = \frac{q_1 + q_2 - q_1 q_2}{q_1(1-q_2)} v_S$$
$$v_{S-i} = \frac{q_1 + q_2 - 2q_1 q_2}{q_1(1-q_2)} v_{S-i+1} - \frac{q_2(1-q_1)}{q_1(1-q_2)} v_{S-i+2}\ , \quad i = 2, \ldots, S-1\ . \tag{A.3}$$

Thus,

$$v_S = \epsilon k_2 v_{S-1} + O(\epsilon^2)$$
$$v_{S-i} = \sum_{j=0}^{i} \left(\frac{q_2(1-q_1) + O(\epsilon^2)}{q_1(1-q_2)}\right) v_{S-1}, \quad i = 2, \ldots S-1\ . \tag{A.4}$$

From (A.2) and (A.4), it follows that

$$v_1 = 1 - \epsilon k_1 Q\left(\frac{k_2}{k_1}, S-1\right) + O(\epsilon^2)\ .$$

Therefore,

$$\begin{aligned} PR_c(k_1, N_1, k_2, N_2, S) &= q_q v_1 \\ &= 1 - \left[k_2 + k_1 Q\left(\frac{k_2}{k_1}, S-1\right)\right]\epsilon + O(\epsilon^2) \\ &= PR_o(k_1, S-1, k_2) + O(\epsilon^2)\ , \end{aligned}$$

$$WIP_c(k_1, N_1, k_2, N_2, S) = \sum_{j=1}^{S-1} v_j + O(\epsilon)$$
$$= \begin{cases} WIP_o(k_1, S-1, k_2) + O(\epsilon) & \text{if } N_1 \le N_2\ , \\ S - WIP_o(k_1, S-1, k_1) + O(\epsilon) & \text{if } N_1 > N_2\ . \end{cases}$$

If $\min(N_1, N_2) < S \le \max(N_1, N_2)$, no starvation of one machine and no blockage of the other machine can occur. Therefore,

$$PR_c(k_1, N_1, k_2, N_2, S) = PR_o(k_1, \min(N_1, N_2), k_2) + O(\epsilon^2)$$
$$WIP_c(k_1, N_1, k_2, N_2) = \begin{cases} WIP_o(k_1, N_1, k_2) + O(\epsilon) & \text{if } N_1 \le N_2\ , \\ S - WIP_o(k_1, N_2, k_1) + O(\epsilon) & \text{if } N_1 > N_2\ . \end{cases}$$

Finally, if $\max(N_1, N_2) \le N_1 + N_2$, the smaller buffer has at least $S - \max(N_1, N_2)$ parts. If $N_1 \le N_2$, then $v_i(n), i = 1, \ldots, N_1$, obey the following equation:

$$v_l(n) = 1, l = 1, \ldots, L,\ L = S - N_2\ ,$$

$$\begin{aligned} v_{L+1}(n+1) &= v_{L+1}(n) + (1 - v_{L+1}(n))q_1 \\ &\quad + [v_{L+2}(n) + (v_{L+1}(n) - v_{L+2}(n))q_1]\, q_2 \\ &\quad - [v_{L+1}(n) + (1 - v_{L+1}(n))q_1]\, q_2\ , \\ v_i(n+1) &= v_i(n) + (v_{i-1}(n) - v_i(n))q_1 \\ &\quad + [v_{i+1}(n) + (v_i(n) - v_{i+1}(n))q_1]\, q_2 \\ &\quad - [v_i(n) + (v_{i-1}(n) - v_i(n))q_1]\, q_2\ , \\ v_{N_1}(n+1) &= v_{N_1}(n) + (v_{N_1-1}(n) - v_{N_1}(n))q_1 \\ &\quad + v_{N_1}(n) q_1 q_2 - [v_{N_1}(n) + (v_{N_1-1}(n) \\ &\quad - v_{N_1}(n))q_1]\, q_2\ . \end{aligned} \tag{A.5}$$

In the steady state, we obtain

$$(1 - v_{N_1})q_1 + v_{N_1} q_1 q_2 - [v_{L+1} + (1 - v_{L+1})q_1]q_2 = 0 \tag{A.6}$$

and

$$v_{N_1-i} = \sum_{j=0}^{i} \left(\frac{q_2(1-q_1)}{q_1(1-q_2)}\right)^j v_{N_i}, \quad i = 0, \ldots, N_1 - L - 1\ . \tag{A.7}$$

Then

$$v_{N_i} = Q(A, N_1 - L) v_{L+1}\ . \tag{A.8}$$

where $A = q_2(1-q_1)/(q_1(1-q_2))$. From (A.6) and (A.8),

$$v_{L+1} = \frac{q_1(1-q_2)}{q_2(1-q_1) + q_1(1-q_2)Q(A, N_1 - L)}\ .$$

Thus, from (A.6)

$$PR_c(k_1, N_1, k_2, N_2, S) = \frac{q_1 q_2 [(1-q_1) + q_1(1-q_2)Q(A, N_1-L)]}{q_2(1-q_1) + q_1(1-q_2)Q(A, N_1 - L)}\ . \tag{A.9}$$

Since $q_i = 1 - \epsilon k_i$ and $Q(\frac{k_1}{k_2}, N_1 - L) = (1 - k_1/k_2)/(1 - (k_1/k_2)^{N_1 - L})$, from (A.9) we find

$$PR_c(k_1, N_1, k_2, N_2, S) = 1 - \left[k_1 + k_2 Q\left(\frac{k_1}{k_2}, N_1 - L + 1\right)\right]\epsilon + O(\epsilon^2)\ .$$

If $N_1 > N_2$, using Theorem 2.6 of Lim *et al.* (1990), an analogous expression is obtained. Thus,

$$PR_c(k_1, N_1, k_2, N_2, S) = PR_o(k_1, N_1 + N_2 - S + 1, k_2) + O(\epsilon^2)\ ,$$

$$WIP_c(k_1, N_1, k_2, N_2, S)$$
$$= \begin{cases} S - N_2 - 1 + \sum_{j=S-N_2}^{N_1} v_j, & \text{if } N_1 \le N_2\ , \\ S - (S - N_1 - 1 + \sum_{j=S-N_1}^{N_2} v_j), & \text{if } N_1 > N_2\ , \end{cases}$$
$$= \begin{cases} S - N_2 - 1 + WIP_o(k_1, N_1 + N_2 - S + 1, k_2) + O(\epsilon), & \text{if } N_1 \le N_2\ , \\ S - (S - N_1 - 1 + WIP_o(k_1, N_1 + N_2 - S + 1, k_1) + O(\epsilon), & \text{if } N_1 > N_2\ , \end{cases}$$