TR diss 2873

Stellingen

behorende bij het proefschrift getiteld:

On Production Lines with Blocking

Auteur:

Erik van Bracht

1. Onze westerse cultuur veronderstelt dat we allen in *dezelfde* wereld leven waarvan ieder individu zijn eigen voorstelling maakt.
   Deze gedachte veroorzaakt een door filosofen veel beschreven verschil tussen het 'wezen', de unieke wereld zoals deze voor iedereen noodzakelijk hetzelfde moet zijn, en de 'schijn', een wereld zoals deze voor ieder individu anders kan zijn.
   In tegenstelling echter tot wat in onze cultuur wordt verondersteld leven individuen in *verschillende* werelden waarin de wereld zoals deze voor ieder individu werkelijk is overeenkomt met de wereld zoals deze voor ieder individu lijkt te zijn:
   Het verschil tussen wezen en schijn is in wezen schijn.

2. Illusie is niet een andere kijk op realiteit, maar de enige manier van leven.

3. Communicatie veronderstelt een niet aantoonbare overeenstemming tussen werkelijke en geprojecteerde gevoelens.

4. Tijdens de Renaissance is de ruggengraat van de westerse samenleving overgegaan van de dogma's van de kerk naar de dogma's van de experimentele wetenschap. Het is echter zeer de vraag of de gangbare dogma's van de experimentele wetenschap de ontwikkeling van de menselijke geest niet *evenveel* in de weg staan als de dogma's van de kerk tijdens de Middeleeuwen.

5. Niemand weet waarom een bal valt.

6. Er zijn zowel echte en onechte sporters als echte en onechte intellectuelen, zij die zichzelf uitdagen en zij die anderen willen overtreffen.

7. Een universiteit is als een klooster, waarin niet novices maar studenten, niet monniken maar wetenschappers zich overgeven aan een geloof, het geloof dat we de wereld nog beter zullen begrijpen als we haar nog beter beschrijven. Echter, de wereld is als de liefde, zelfs een perfecte beschrijving mist volledig de kern.

8. Een absolute betekenis van de begrippen 'waar' en 'onwaar' en de begrippen 'goed' en 'slecht' zal nooit objectief vastgesteld kunnen worden.

9. De motivatie achter het bedrijven van wiskunde is dezelfde als die achter de liefde: het verlangen naar een volmaakte toestand waarin werkelijkheid en vanzelfsprekendheid fuseren.

TR diss
2873

On Production Lines with Blocking

688900
319514
TR diss 2873

On Production Lines with Blocking

PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus,
Prof. ir. K.F. Wakker,
in het openbaar te verdedigen
ten overstaan van een commissie
aangewezen door het College van Dekanen
op vrijdag 20 december 1996 te 10.30 uur

door

Erik VAN BRACHT

wiskundig ingenieur,
geboren te Gouda.

BIBLIOTHEEK TECHNISCHE UNIVERSITEIT
Prometheusplein 1
2628 ZC
DELFT

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. G.J. Olsder

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. G.J. Olsder	TU Delft, NL, promotor
Dr. J.W. van der Woude	TU Delft, NL
Prof. F. Baccelli	INRIA, Sophia Antipolis, FR
Prof. dr. ir. A.C.P.M Backx	TU Eindhoven, NL
Prof. dr. ir. O.J. Boxma	CWI, Amsterdam, NL
Prof. dr. R.M. Cooke	TU Delft, NL
Prof. dr. ir. A.W. Heemink	TU Delft, NL

Van Bracht, Erik

On Production Lines with Blocking / Erik van Bracht.
Thesis Technische Universiteit Delft.
With summary in Dutch.
ISBN 90-9009846-1
NUGI 811
Subject headings: queueing networks with blocking / decomposition /
Markov chains

The author allows everyone who is interested in parts from this thesis to make as many reproductions as he likes provided that the source of those reproductions, regardless of the kind of information storage and retrieval system that is used, can be traced by everyone who owns them.

Preface

This thesis is the result of five years research in the Systems Theory Group of the section of Applied Analysis at the department of Technical Mathematics and Computer Science of the Delft University of Technology in the Netherlands.

I am the result of twenty-eight years of support from family and friends and education in schools. My education started in Gouda with two years nursery school, six years primary school followed by six years secondary school. Next, for reasons that are a combination of interest in puzzles and talent, I chose to study Applied Mathematics and be a student for five years at the Delft University of Technology. Having finished this period successfully, I was asked by Prof. dr. G.J. Olsder to continue studies for another four years as a PhD student in an academic research project team for optimization of a production line of a well-known lamp manufacturer. After some hesitation, for reasons that are a combination of pride, unawareness of other possibilities, the feeling that I would never have a similar second offer and my endless interest in mathematical puzzles, I decided to take this opportunity.

The personal consequences of this decision for me were big during the next five years. I think that because of my background, growing up in a protected surrounding, and my talent for mathematics at school, until the time of the decision I never needed to put mental effort in reaching goals. Until then I was used to putting mental effort in things only because of the mental pressure of people who expected me to do so. As a PhD student, however, there appeared to be practically no one with expectations about the results of my work and there appeared to be very few people around who could assist or with whom I could communicate about the fundamental problems that I met. I think that this was partly due to my own attitude: I could have searched better in the existing literature and I could have searched for and met more people around this planet who are busy with similar problems. However, this was not my style. I have tried to tackle the fundamental problems in my own personal way.

The most difficult problems that I met, however, were *not* in the field of mathematics, but in the field of my personal philosophy, psychology and management. I doubted the purpose of science and had problems finding the corresponding personal goals and motivation to reach them. One of the main problems that I met and that I would like to discuss here is

of a pure philosophic nature.

I think that, unconsciously, my interest for mathematics came from the urge to gain insight into the phenomena of the world around me. It is only now that I see clearly that mathematics is a nice tool to *describe* phenomena of this world, but will never be able to reveal the answer to the more interesting question "*Why ?*".

It is maybe because of too high expectations that I feel disappointed in the answers that science offers to mankind. In my view science only records and simplifies phenomena, but does *never* explain. Take for instance the reasoning that objects fall on earth *because of* the gravitational field and Newton's laws. I find that this *abuse of reasoning* is characteristic of the scientific method. Strictly spoken, the only thing that Newton's laws and the gravitational field can do for us is accurately describe observations of falling objects in the past. Most scientists are satisfied with Newton's description and behave as if this were an explanation. But they do not see that Newton's description does *not* relate to the answer to the question: "*Why* did objects fall towards earth?". The scientific answer to the last question is not deeper than "Because objects have always fallen towards earth." No matter how accurately we can describe objects that fell in the past, is this really an answer to our question? Should we call this an explanation? My personal answer to the last question is clearly "No!".

Description of phenomena *can* be seen as a first step towards understanding. The problem is that after the first scientific step of description already we are facing the frontiers of pure science. A reasoning that is still contained within this first step could be "*If* future objects in the year 2000 fall in a similar way as we have observed in the past *then* objects in the year 2000 will fall towards earth." A next step however, I think unconsciously taken by almost all scientists, is almost necessarily a step *over* the borders of science into religion. I think for instance that it should be seen as *religion* to remove the condition and reason that "Objects in the year 2000 will fall towards earth *because* all observed objects in the past have done so" (Which is what we basically do if we *explain* falling objects in the year 2000 by means of gravitational fields and Newton's laws.) No one has observed falling objects in the year 2000, but still we *believe* that these future observations will be similar to the experiences that we have with falling objects in the past: we *feel* confident that in the year 2000 objects will fall towards earth as well. This *religion* is a subtle but deeply rooted unconscious *belief* of mankind in the necessity of the continuity of the world in which we live. This belief is natural because everyday life would become impossible if we doubted the direction of falling objects, but that does not contradict my fundamental point here. In this sense, I am convinced that science itself is in fact a pre-eminent religion. Science is a religion that *claims* the necessity of unobserved phenomena based on the continuation of observed ones.

A most interesting question to me is what the world would look like if we refused this

dogma and refused the way in which science deals with the necessity of continuation of observed phenomena. If we refuse this dogma we have an infinite number of possible descriptions based on the same set of observations of which I will distinguish three different types:

- **Rejection of continuation in the future (Extrapolation to the future)**
  Science tells us only for reasons of continuity that objects will fall towards earth in the year 2000. If we drop the dogma of continuation we can believe that objects will continue falling towards earth in accordance with past experiences *until* the year 2000 and then stop doing so. We can believe that, because of a sort of apocalypse, objects will suddenly start falling towards the sky from the year 2000 on.

- **Rejection of continuation between observations (Interpolation)**
  Refusing the dogma of continuation, we can even refuse conclusions from continuation *between* the observed phenomena. Science tells us for instance that objects have *always* fallen towards earth. This includes a claim on the past behaviour of all unobserved falling objects as well. If we refuse the dogma of continuation we can believe for instance that on a day in the year 1900, even if somewhere on an island near the equator people have observed hundreds of coconuts falling towards the earth, still one unobserved coconut did not fall towards earth but fell towards the sky. In this way we may believe for instance that there are or might have been a lot of unobserved objects falling towards the sky *between* the observed ones that fell towards the earth. There could even be a causal relation between the fact of the observation itself and the behaviour that falling objects have: maybe objects fall towards earth if and only if they can be observed!

- **Rejection of continuation in the past (Extrapolation to the past)**
  An interesting and classic example that refuses the dogma of continuation is taken from the Bible. We can join some (other) religious people and believe that the earth has been created by a "higher power" about 40.000 years ago. The consequences of this creation is that obviously before 40.000 years ago objects did not fall towards earth because they did not even exist. By a discontinuous act of creation suddenly there was an earth with objects in it. These objects were given the property to sustain falling towards earth just as we have observed during the last few thousands of years.

I think that most scientists do not take these solutions as serious alternatives, but they forget that they do not have better arguments, because as long as solutions are in accordance with past observations they just cannot be rejected.

My conclusion is that an infinite number of different continuous and discontinuous descriptions exist that match with all past observations! Each of these descriptions are *as likely as* the others! Science has no fundamental arguments other than a *belief* in continuity to

claim a best. (I use the adjective *likely* on purpose because it is often used as a counter-argument in this context but has in fact no fundaments other than expressing the same feeling of continuity in a different way.)

However, if we have a closer look at the arguments of science, we see that they are even weaker than I have just described since even if we *accept* to believe in the continuity of phenomena in our world then we are still facing a fundamental problem of non-uniqueness. My point here is that there are also an infinite number of different *continuous* descriptions that match with all observations from the past, just as there exist an infinite number of continuous functions that match with a function of which we do not know the behaviour on an interval. The solution that science finds for this problem is an additional *belief* in the validity of the most "elegant" description: if a shorter continuous description can describe more observed phenomena then scientists *believe* that it will necessarily give *better explanations*. Again there is no other argument than a *belief* in the necessity of the shortest and simplest description.

In order to overcome all these fundamental problems of science, science itself comes up with a solution by means of a new concept denoted by *mathematical model*. The concept of the mathematical model is the solution that should bridge the philosophic gap between the descriptions by means of mathematics and reality (whatever that may be). But, even though the concept of the mathematical models certainly makes the description less pretentious, it does not change the religion of science. It does not stop scientists from believing in the necessity of a continuous world and it does not stop scientists from believing in the necessity of the shortest and simplest descriptions.

I think that science does serve a fundamental service to mankind, a service that in my view can basically be described by the following *psychological* effects that it has:

- the satisfaction of the need to feel a basic structure.
- the satisfaction of the need to improve in life.
- the satisfaction of the need to control phenomena in order to maintain society.
- the satisfaction of the vanity and curiosity of scientists.
- the reassurance against fear for the future.

However, no matter how well scientists have been able to predict and control phenomena, it is *my belief* that science will always be unable to find a unique description just as mankind will always be unable to reveal the answer to the question "*Why ?*".

The purpose for me of finding solutions to mathematical puzzles therefore is not anymore the satisfaction of gaining insight in the phenomena of our world. Mathematics in my

present view does *not* relate to explanations of phenomena in this world. The purpose of applied mathematics in science is only to give mankind a *psychological service* by means of reassurance with structures, improvements and predictions of phenomena. It is not so important whether these structures, improvements and predictions come from real explanations or from a *belief* in a best choice taken from an infinite number of possible structures. The only thing that really matters is that they have the reassuring psychological effects.

In this way I lost my believe in the concept of understanding better "reality" by means of mathematical descriptions. Reality in my present view is a *personal* perception of a world based on a *personal* set of past and present experiences and therefore it *cannot be* a unique entity. Each of these personal realities is as valuable, as good or bad, as likely and as true or false as the other. Religions, in my view, try to unify personal realities into a collective reality. This unification certainly serves the psychological and social needs in a stable society, but does not make a collective reality better, more likely, or more true than any personal one. If someone has a perception of reality in which tomorrow this thesis will fall towards the sky then this perception is as good as, as likely as, and as true as a perception of scientists in which it will fall towards the earth.

I hope that this personal perception of science will not demotivate the reader as it once demotivated myself. It is obvious that *every structure* of the world around us starts necessarily in a belief and that a stable society just needs a clear structure. Moreover, science has facilitated life and increased the expectations of lifetime. It controls and predicts phenomena better than ever before. Even though this doesn't change my mind on basic philosophic points, it certainly offers the opportunity to reach goals and satisfy people's needs. One of my personal goals, that I seem to have reached here, is the construction of a PhD thesis. I admit that I am proud of it.

EvB

Delft
August 1996

Contents

Chapter 1

Introduction

1.1 The Scope of this Thesis

This thesis is the result of five years study on the subject of production lines with blocking. The main purpose of the research was to obtain a computable mathematical model that describes a real production line in a car lamp factory.

If we have a first look at the car lamp production line that we want to describe then we distinguish machines that manipulate products and product parts that *wait* for their manipulations, just like a post office where we can distinguish counters with receptionists and customers who *wait* for service in lines in front of them. The receptionists *serve* customers just like the machines *serve* product parts with manipulations. The receptionists and the machines can therefore be denoted by the more general term *server*. The lines of customers or product parts in front of the servers can be denoted by a the general term *queue*. In other words, in order to describe the car lamp production line, we have to find the most suitable description of a well-known phenomenon called *waiting*. This is why we enter a huge field in probability theory that is involved in finding mathematical descriptions of objects that are waiting in lines. This field in probability theory is called *queueing theory*.

1.1.1 The Level of the Mathematics in this Thesis

In this thesis we assume that the reader has some basic knowledge of mathematical probability theory and stochastic processes. For an introduction into basic mathematical probability theory there exist an enormous amount of books. Between those, for the unsuspecting layman who has the courage to start at the lowest level, we recommend to use [HL64] or [Blo89]. Apart from knowledge of basic probability theory it is useful to have some knowledge about stochastic processes. Books that we recommend because they take the step from probability towards stochastic processes in a most comprehensible manner are [GS87],

[Goo88] and [All90]. In fact in this thesis we aim at readers who do not have a higher level than required for these last three references. Therefore, the knowledge of stochastic processes that one needs in order to read this thesis is minimal. This is also because we tried to increase the accessibility of this thesis by the inclusion of a long section in the introduction that contains a description of the most simple class of stochastic processes: the homogeneous discrete-time finite Markov chains. It is only in the last chapter, Chapther 6, that we make a small step beyond this class and use more complicated stochastic processes. In a lot of proofs in this thesis, of which most are mentioned in Appendix A and Appendix B, we assume a sufficient level in basic calculus.

1.1.2 Zooming Into the Field

The field of queueing theory is so extensive that we can by no means give a clear overview. Therefore it is wise to examine the properties of the car lamp production line in order to zoom into smaller specialized fields. If we have a closer look at the car lamp production line we observe

- *several servers* and *several queues* that are linked in such a way that product parts that have been served by machines are put in next queues in front other machines. This means that we are not dealing with a single server (as extensively described in [Coh69], and for Markovian queues in [Sha90]), but with a more complicated *queueing network* of servers.

- raw product parts that enter into the queueing network *from outside*, and finished lamps that leave the network *into the outside world.* This implies that the car lamp production line is an *open* queueing network.

- queues that have a *finite capacity.* This means that queues can become *full* which introduces one of the most complicated phenomenon in queueing networks: *blocking.*

- only one machine that takes product parts from two different queues. This implies that the car lamp production line is not a serial queueing network, but a queueing network that has a configuration with merging.

- short term failures that occur with the manipulations on product parts and long term failures that occur if machines break down. These breakdowns make that the machines are *unreliable.* This requires special properties of the service time distributions.

- machines that have fixed cycle times, which is also denoted by the term "indexing machines". At the end of each cycle exactly one product part comes from a machine.

We conclude that we have to focus on *open queueing networks with blocking*, which is still a big area, but which is small enough to obtain a reasonable overview.

1.1.3 Methods to Solve Queueing Networks with Blocking

If we examine the literature in this area, for this we refer for instance to [PA89], then we come to the conclusion that it deals with one of the hardest problems in queueing theory. The problem with blocking is that it induces complex correlations between the behaviour of two subsequent servers. As far as we know there has not yet been found an elegant solution by means of which these complex correlations can be described. All methods in this area are focused on finding better numerical approximations and improvement of theoretical upper and lower bounds for the stationary behaviour of such networks. What makes this area of research so interesting is that one has the feeling that the mechanisms of its queueing systems are simple, common and natural but each attempt to solve for simple and exact analytic solutions ends up in a failure.

There are some directions in which researchers have tried to find solutions. We think that the mainstreams of these directions are the following.

- The direction of the *Perturbation Analysis*, which in fact is a technique that estimates sensitivities and gradients from real time measurements in the network. For this method we refer to [HC91], [HET79] and [Cas93].

- The direction of the *Product Form Solutions*, which assumes that the stationary distribution can (approximately) be written in the special product form. This product form turns out to give exact solutions in a well-defined area of non-blocking networks. For these product form solutions we refer to [Dij93], [Wal88] and [Dij89].

- The direction of the *Petri nets* and *Timed Event Graphs*, for which we refer to [BCOQ92]. In this direction we obtain very "elegant" descriptions of queueing networks with blocking. It even gives nice expressions in the setting of a new algebra that can be used for queues with deterministic service times. However, the expressions that we obtain in this direction for queues with stochastic service times are still very complicated, so complicated that solving the exact stationary distributions analytically or numerically is still beyond our reach. Even though it is yet not possible to apply the methods in this direction for stochastic queueing networks, we think that this direction should be seen as the hardest but most promising one. In the last chapter in this thesis we will make a small step into this direction.

- The direction of *Decomposition Techniques* in queueing networks. This is the most classic direction that tries to project the behaviour of queues in the network on the behaviour of a predefined class of single server queues. There is a large amount of methods in this direction. For an extensive overview of methods in this direction we refer to [DG92].

Of course we can think of other directions and directions that consist of combinations of the ones that we just described. We just consider these four directions as the leading

ones at this moment. The last direction, the direction of *Decomposition Techniques*, is the direction that we chose in order to find the stationary behaviour of the car lamp production line in this thesis. We refer to [MT90], [HB67], [DRXL88] and [Ger93] in particular for descriptions of methods for queues with blocking in series with "unreliable" servers. Similar serial queueing networks with "unreliable" servers will also be treated in this thesis. These articles therefore address similar problems. A fundamental difference of the methods in this thesis with the methods derived in those articles, is the *discrete-time* nature of the models here versus the continuous time nature of similar models in [HB67], [DRXL88] and [Ger89]. We have chosen for the less usual *discrete time* models in (the main part of) this thesis because they fit better with the discrete time nature of the "indexing machines" in the car lamp production line.

It is obvious that the modest list of the author's own publications is related to this thesis as well, see [BOW91, Bra91, Bra93, Bra94, Bra95].

1.1.4 An Outline of this Thesis

This thesis is divided in two parts. The first big part contains Chapter 1 up to Chapter 5. The second small part consists of Chapter 6 only.

The target of the first part is to find a suitable description by means of a mathematical model of a real car lamp production line by means of the most simple subclass of Markov chains, the homogeneous discrete-time finite Markov chains. We use *discrete-time* finite Markov chains in order to obtain a better model of the discrete behaviour of the "indexing machines" in the car lamp production line. These "indexing machines" have *fixed* cycle times. Although the discrete-time Markov chains can model better this discrete-time behaviour, we have to cope with the problem that they model machines with the same cycle time. An essential difference with most other models for such lines that we found in the literature is that most of them use continuous- time stochastic processes.
This first big part can be divided into three smaller parts:

- Chapter 1 that contains this section about the scope of this thesis and a rather long introduction into the principles that we use for our mathematical models, such as a description of the notions production line and discrete-time Markov chain.

- Chapter 2, Chapter 3 and Chapter 4 in which we describe how we use decomposition methods in combination of the principles from Chapter 1 in order to derive computable mathematical models for *serial* production lines. The first chapter in this part, Chapter 2, describes how we deal with short term failures (geometric service time distributions) in serial production lines. The second chapter in this part, Chapter 3, describes how we deal with long term failures (breakdown) in serial production lines. The third and last chapter in this part, Chapter 4, describes how we can combine the ideas from the previous two chapters and describes therefore how we deal

with a combination of short term (geometric service times) and long term failures (breakdown) in such lines.

- Chapter 5 that describes how we extend the decomposition methods from the *serial* configurations in previous chapters and apply them to the real *non-serial* car lamp production line.

In a way we have never been satisfied with the solutions and methods that come from the mathematical models in the first part. We felt dissatisfied for reasons that are a combination of a feeling of a lack of "elegance" and awareness of the limitations of the descriptions in the first part. Therefore we have spent *most* of our research in a completely different direction in spite of the fact that it most probably would not lead to a possible application in the car lamp production line. This other direction that deals with stochastic timed event graphs is decribed in Chapter 6. We believe that in this direction the most interesting new results can be found. A lot of current research is already in this field at the moment.

In this last part in Chapter 6 we introduce a new model for serial queueing networks with blocking based on stochastic timed event graphs. We believe that the concept in this model in which we do not make a fundamental difference between the positions in a queue and the positions in servers can be seen as a small breakthrough. Therefore we believe also that reasoning along the lines of this concept deserves more research, because it is likely to offer new methods to approximate or maybe even find exact waiting time distributions in queueing networks with blocking.

1.2 Production Lines

In the previous section we described the scope of this thesis and argued that the scope of the research is the area of queueing networks with blocking. Because in this thesis we are dealing with manufacturing systems only we will not use the notion *open queueing network with blocking* but from now on we will use the equivalent notion *production line* instead. Therefore we will not use the terms *queue* and *server*, but from now on we will use the equivalent terms *buffer* and *machine* instead.

Before we examine production lines and the special possible properties we have to describe what we exactly mean here with the notion *production line* in general. We will describe this notion accurately in the next subsection.

1.2.1 Configuration

Whenever we use the notion *production line* we refer to a *directed graph* that consists of a set of items called *machines* denoted by $\mathcal{M}$, a set of items called *buffers* denoted by $\mathcal{B}$

and a set of *arcs* denoted by $\mathcal{A} \subset (\mathcal{M} \times \mathcal{B}) \cup (\mathcal{B} \times \mathcal{M})$ such that $\forall M \in \mathcal{M},\ \exists B_{\text{in}}, B_{\text{out}} \in \mathcal{B}$ such that $(B_{\text{in}}, M) \in \mathcal{A}$ and $(M, B_{\text{out}}) \in \mathcal{A}$.
In other words a production line is a collection of machines and buffers connected by arcs. Buffers and machines are never directly connected among themselves. Machines always have both input and output arcs.

Example.

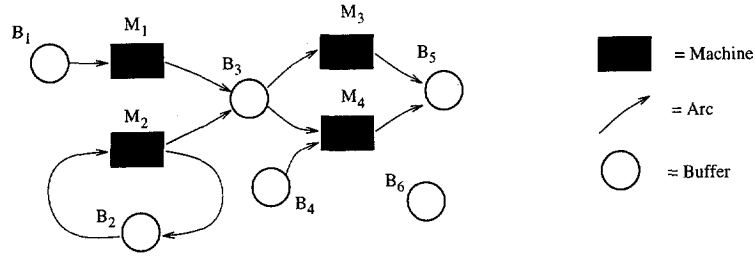

Figure 1.1: An arbitrary production line.

A picture of an arbitrary production line is shown in Figure 1.1.
In this example

$$\begin{aligned}
\mathcal{M} &= \{M_1,\ M_2,\ M_3,\ M_4\},\\
\mathcal{B} &= \{B_1,\ B_2,\ B_3,\ B_4,\ B_5,\ B_6\},\\
\mathcal{A} &= \{(B_1, M_1),\ (B_2, M_2),\ (B_3, M_3),\ (B_3, M_4),\ (B_4, M_4)\} \cup \ldots\\
&\quad \{(M_1, B_3),\ (M_2, B_2),\ (M_2, B_3),\ (M_3, B_5),\ (M_4, B_5)\}.
\end{aligned}$$

□

If $(x, y) \in \mathcal{A}$ then we say that x is *upstream* with respect to y and y is *downstream* with respect to x. In the example of Figure 1.1 machine M_4 has two upstream buffers B_3 and B_4 and buffer B_3 has two downstream machines M_3 and M_4.

A buffer without an upstream machine is denoted by the term *input buffer* or *source*, a buffer without a downstream machine is denoted by the term *output buffer* or *sink*. The example of Figure 1.1 has three input buffers B_1, B_4, B_6 and it has two output buffers B_5, B_6.

1.2.2 Products and Product Parts

The notion "production line" makes us think that something is produced. Of course the purpose of a production line is to produce certain things called *products* that can be sold to customers in order to satisfy their needs and to make profit. So far the production line considered as a directed graph is a static object, but if we zoom inside the line we can see discrete parts moving from machines to buffers and from buffers to machines following the directions of the arcs. We put *raw parts* into the input buffers, then the *product parts* go through the production line following the arcs, are disconnected, manipulated or reconnected each time they visit a machine in such a way that after some time they may end up as finished *products* as soon as they reach an output buffer.
It is assumed that the products and product parts are of a *discrete* nature, that means that they cannot be handled in any amount other than natural numbers. (This is in contrast with "liquid" product parts for instance in a paint factory.)

1.2.3 Machines

As we explained before a production line consists partly of machines. Machines are entities that are supposed to work on product parts. They are supposed to repeat the following three actions continuously:

- Examine if all *upstream* buffers have a product part. If so then take one new product part from all downstream buffers.
- Manipulate the product parts if taken.
- Examine if all *downstream* buffers have space left for a manipulated product part. If so then put one manipulated product part in all upstream buffers.

The procedure of the three actions mentioned above we will denote as a *cycle*. Each machine continuously repeats cycles with a certain speed, the time in which a machine completes one cycle is denoted as *cycle time*.

Machine failures

The cycles as described in the previous paragraph consist of actions that machines are supposed to do continuously. Unfortunately machines are prone to failure! When a major failure occurs such that the machine is not able to continue manipulating new product parts we talk about a *machine failure*. When a machine failure occurs the machine stops its cycles and waits until an operator comes and fixes the problem. We say that the machine is in *down* state. After having fixed the problem the operator restarts the machine and the machine resumes its cycles. We say that the machine is in *up* state again. The transition from a machine's up to down state is denoted as *breakdown*. The transition from a machine's down to up state is denoted as *repair*.

Product part failures

Within one cycle a machine manipulates one product part. Of course failures of manipulations performed on the product parts can occur also: a product part can break during the manipulations or a connection made between product parts is not made properly. These product parts that fail manipulations are often useless for next manipulations in the production process since they will never result in a product that fulfils the requirements of customers. If such a *product part failure* occurs and is detected by the machine, then there are two possibilities: either the product part is put back in the upstream buffer in order to correct for the failure the next time that it will be taken, or the broken product part will be totally removed from the production process because the failure cannot be corrected. If the broken part will be removed from the production process then it will be put in a *garbage bin* just after manipulation by the machine. The latter is denoted by the term *scrapping.*

Notice that the difference between machine failures and product part failures is such that when a product failure occurs the machine continues its cycles and when a machine failure occurs machine cycles stop.

1.2.4 Buffers

Products and product parts in the production line are assumed to be stored in a buffer whenever they are not under manipulation of a machine. Buffers are nothing more than spaces for storage of products and product parts. Unfortunately space for storage is naturally limited in almost every realistic situation. The maximal number of products that a specific buffer can contain is denoted by the term *capacity.* The capacity of a buffer is always a positive natural number.

It is assumed in this thesis, except for the models in the last chapter, that buffers don't have transportation times.

Buffers have a dynamically changing number of product parts in it. This implies that buffers can become empty and full. When a buffer is empty and a machine tries to take a part from it, we speak of *starvation.* When a buffer is full and a machine tries to put a part in it we speak of *blocking.* Blocking and starvation are important notions for production lines in which buffers have limited capacities.

1.3 Introduction to Markov Chains

In this section the notion of *Markov chain* will be described and basic tools dealing with these chains will be explained. For a more thorough explanation and description of Markov chains we refer to [GS87], [Goo88] and [All90].

A random process is a family of random variables $\{X(t) : t \in T\}$ indexed by some set T. To make life easier we will assume that the random variables are discrete such that $X(t)$ takes values over a finite set $\mathcal{S} = \{0, 1, .., N\}$ of states. Moreover, the set T consists of natural numbers only. This makes the random process a discrete time stochastic process. Amongst the various types of stochastic processes we can distinguish a type that has proved to be very useful for modelling real life systems: the Markov chain.

1.3.1 The Markov Chain

The very basic property of this type of processes is that it is memory-less in the following sense:

$$\mathbb{P}\mathrm{r}(X(t) = s|X(0), X(1), .., X(t-1)) = \mathbb{P}\mathrm{r}(X(t) = s|X(t-1)) \ . \tag{1.1}$$

When a stochastic process has the above property, it can be fully characterized by the so called *transition probabilities* $\mathbb{P}\mathrm{r}(X(t+1) = j|X(t) = i)$. In general the transition probabilities depend on i,j and t. For the sake of simplicity we restrict ourselves to the time independent homogeneous Markov chains, which means that the dependency on t is dropped. This restriction brings us to to class of *homogeneous discrete-time finite Markov chains*, in which the random process can be completely characterized by the *transition matrix* $P = (p_{ij})$ consisting of the transition probabilities

$$p_{i,j} = p_{ij} = \mathbb{P}\mathrm{r}(X(t+1) = i|X(t) = j). \tag{1.2}$$

In the following the comma is omitted in case it cannot cause confusion. This transition matrix P belongs to the class of *stochastic matrices*, that means that P has the following properties:

(a) P has non-negative entries, equivalently $p_{ij} \geq 0$,
(b) P has column sums equal to one, equivalently $\sum_i p_{ij} = 1$.

To even refine the class of Markov chains more, we will deal only with the case where all states are *ergodic*, which means that they are *aperiodic* and *non-null persistent*, such that the set $\mathcal{S}$ of states is *irreducible*. We will not explain the exact meaning of these notions from Markov chain theory here. For our class of homogeneous discrete-time finite Markov chains the *non-null persistent* requirement is always satisfied if the chain is *irreducible*. In order to check if a Markov chain in this class is irreducible do the following: draw a graph that corresponds to the transition matrix such that the nodes are the states in $\mathcal{S}$ and the directed arcs correspond to the non-zero entries in the transition matrix P. The Markov chain is irreducible if and only if it is possible to go from any node to any other node by means of a sequence of directed arcs. Such a graph is called a *strongly connected graph* in

graph theory. A state in the graph is *aperiodic* if the number of arcs needed to go from that state and return to the same state is not "necessarily" a multiple of a fixed prime number. A-periodicity is assured if each node has an output arc directed to itself, which is the case in almost all Markov chains in this thesis.

At each time t we have a random variable $X(t)$ which takes its values from a finite set $\mathcal{S} = \{0, 1, .., N\}$. Corresponding to each time $t \in \{0, 1, 2, ..\}$ there is a *distribution vector* or briefly *distribution* $\pi(t) = (\pi_i(t) : i \in \mathcal{S})$ such that $\pi_i(t) = \mathbb{P}\mathrm{r}(X(t) = i)$. In the class of homogeneous discrete-time finite Markov chains the time evolution of the distribution vector can easily described by the following matrix multiplication:

$$\pi(t+1) = P\pi(t) \ ,$$

and, as a consequence:

$$\pi(t) = P^t \pi(0) \ . \tag{1.3}$$

So, given an initial distribution $\pi(0)$, the time evolution of the distributions is fixed by means of the previous formulas.
A next question is what happens if time evolves: when time evolves does the distribution vector converge to a limiting distribution or not? In the class of Markov chains that we just described it has been proved that there always exists a *stationary distribution* vector $\pi = (\pi_i : i \in \mathcal{S})$ such that:

$$\begin{aligned} \lim_{t\to\infty} \pi(t) \quad = \quad \pi \quad &= \quad P\,\pi \ , \\ \pi_i \quad &> \quad 0 \ , \\ \sum_i \pi_i \quad &= \quad 1 \ . \end{aligned} \tag{1.4}$$

Remark Note that throughout this thesis we shall use distributions in the form of *column* vectors. This means that for the description of a transition in a Markov process we shall use a stochastic matrix and multiply it from the *right* with a *column* distribution vector. This is not in accordance with the usual conventions in which distributions are *row* vectors and multiplication is from the *left*.

A simple example

In the case of the example shown in Figure 1.2, $\mathcal{S} = \{0, 1, 2\}$ and

$$P = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix}.$$

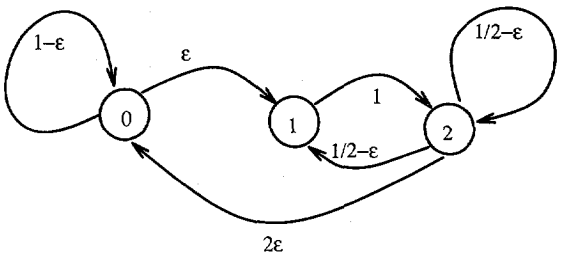

Figure 1.2: Graph of a homogeneous discrete-time finite Markov chain.

Suppose that we start in state 0 "for sure" at time instant 0. That means that the initial condition is as follows:

$$\begin{aligned} \pi_0(0) &= \mathbb{P}\mathrm{r}(X(0)=0) = 1 \,, \\ \pi_1(0) &= \mathbb{P}\mathrm{r}(X(0)=1) = 0 \,, \\ \pi_2(0) &= \mathbb{P}\mathrm{r}(X(0)=2) = 0 \,, \end{aligned}$$

which gives the distribution vector:

$$\pi(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

When time evolves we go from random variable $X(0)$ to $X(1)$. At time 0 we were "sure" to be in state 0, but as one can see from the graph, we can go to either state 0 again or to state 1 from state 0. The probability of a transition to state 0 again has probability $1-\varepsilon$, the probability of a transition to state 1 is equal to ε. That means that after one transition it is not known in which state the system is: it is either in state 0 again with probability $1-\varepsilon$ or in state 1 with probability ε. So now we get:

$$\begin{aligned} \pi_0(1) &= \mathbb{P}\mathrm{r}(X(1)=0) = 1-\varepsilon \,, \\ \pi_1(1) &= \mathbb{P}\mathrm{r}(X(1)=1) = \varepsilon \,, \\ \pi_2(1) &= \mathbb{P}\mathrm{r}(X(1)=2) = 0 \,, \end{aligned}$$

which gives the distribution vector:

$$\pi(1) = \begin{pmatrix} 1-\varepsilon \\ \varepsilon \\ 0 \end{pmatrix} = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = P\,\pi(0) \,.$$

Now we can easily continue like this:

$$\pi(2) \;=\; P\,\pi(1) = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix} \begin{pmatrix} 1-\varepsilon \\ \varepsilon \\ 0 \end{pmatrix} = \begin{pmatrix} (1-\varepsilon)^2 \\ \varepsilon(1-\varepsilon) \\ \varepsilon \end{pmatrix},$$

$$\pi(3) \;=\; P\,\pi(2) = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix} \begin{pmatrix} (1-\varepsilon)^2 \\ \varepsilon(1-\varepsilon) \\ \varepsilon \end{pmatrix} = \begin{pmatrix} (1-\varepsilon)^3 + 2\varepsilon^2 \\ \varepsilon(1-\varepsilon)^2 + \varepsilon(\frac{1}{2}-\varepsilon) \\ \varepsilon(1-\varepsilon) + \varepsilon(\frac{1}{2}-\varepsilon) \end{pmatrix},$$

etcetera.

Since the corresponding graph is strongly connected (you can go from any node to any other node by a sequence of arcs.), there must exist a stationary distribution vector π which has strictly positive entries. To derive this stationary distribution we solve the following equation:

$$\pi = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix} \pi\,,$$

subject to the condition that the sum of the entries of π is equal to 1. The stationary distribution for this example is:

$$\pi = \frac{1}{7+2\varepsilon} \begin{pmatrix} 4 \\ 1+2\varepsilon \\ 2 \end{pmatrix}. \tag{1.5}$$

□

1.3.2 Construction of Clusters

One of the biggest problems when modelling real world by means of Markov chains is that the number of states in the state set $\mathcal{S}$ increases rapidly as the model tends to be more realistic. A natural thing then to do is make groups of states acting as if they were just one state on their own. This *clustering* principle can be applied successfully, as we shall see in this thesis. How one can apply clustering will be described now.

Suppose we have a Markov chain with the (arbitrary) disjoint partitioning of the set $\mathcal{S}$ together with a transition matrix P:

$$\begin{aligned} \mathcal{S} \;&=\; \mathcal{S}^0 + \mathcal{S}^1 + .. + \mathcal{S}^n\,, \\ P \;&=\; \begin{pmatrix} P^{(0,0)} & P^{(0,1)} & . & . & P^{(0,n)} \\ P^{(1,0)} & P^{(1,1)} & . & . & P^{(1,n)} \\ . & . & & & . \\ . & . & & & . \\ P^{(n,0)} & P^{(n,1)} & . & . & P^{(n,n)} \end{pmatrix}, \end{aligned} \tag{1.6}$$

such that the partitioning in states corresponds to the partitioning in the transition matrix P. Suppose we are also given a current distribution π:

$$\pi = \begin{pmatrix} \pi^{(0)} \\ \pi^{(1)} \\ \cdot \\ \cdot \\ \pi^{(n)} \end{pmatrix}, \tag{1.7}$$

whose partitioning again corresponds to the partitioning in the states. Then we can try to setup a new state space $\tilde{\mathcal{S}} = \{0, 1, .., n\}$ with a new transition matrix $\tilde{P}(\pi)$, such that each state i in $\tilde{\mathcal{S}}$ corresponds to the cluster of states $\mathcal{S}^i$ of the original system. If we define for $x \in \mathbb{R}^m$:

$$\| x \|_1 \stackrel{\text{def}}{=} \sum_{i=1}^{m} |x_i|, \tag{1.8}$$

we can formulate the following new clustered Markov chain structure:

$$\begin{aligned} \tilde{\mathcal{S}} &= \{0, 1, .., n\} , \\ \tilde{P}(\pi) &= (\tilde{p}_{ij}(\pi)) , \\ \tilde{p}_{ij}(\pi) &= \frac{\| P^{(i,j)}\pi^{(j)} \|_1}{\| \pi^{(j)} \|_1} , \\ \tilde{\pi}(\pi) &= \begin{pmatrix} \| \pi^{(0)} \|_1 \\ \| \pi^{(1)} \|_1 \\ \cdot \\ \cdot \\ \| \pi^{(n)} \|_1 \end{pmatrix} , \end{aligned} \tag{1.9}$$

Note that the transition probabilities for the clustered version from the current state to the next state are dependent on the current distribution of the non-clustered state! To emphasize this dependency the π is added as an argument for the clustered transition matrix $\tilde{P}(\pi)$. In this way,

$$\pi(t+1) = P \ \ \pi(t) , \tag{1.10}$$

corresponds directly to:

$$\tilde{\pi}(\pi(t+1)) = \tilde{P}(\pi(t)) \ \ \tilde{\pi}(\pi(t)) . \tag{1.11}$$

With respect to stationary distributions we will drop the distribution π as an argument:

$$\begin{aligned} \tilde{\pi} &\stackrel{\text{def}}{=} \lim_{t\to\infty} \tilde{\pi}(\pi(t)), \\ \tilde{P} &\stackrel{\text{def}}{=} \lim_{t\to\infty} \tilde{P}(\pi(t)). \end{aligned} \tag{1.12}$$

Clustering and its notation for partitionings

The clustered version of a Markov chain of course depends on the partitioning chosen in 1.6. For the sake of convenient notation in this thesis we will define first the set $\mathcal{P}(\mathcal{S})$ of all disjunct partitionings of set $\mathcal{S}$ in clusters as:

$$\mathcal{P}(\mathcal{S}) \stackrel{\text{def}}{=} \{ \ (\mathcal{S}_0, \mathcal{S}_1, .., \mathcal{S}_n) \ \mid \ \sum_{i=0}^{n} \mathcal{S}_i = \mathcal{S} \ \} \tag{1.13}$$

Often we want to be clear which partitioning is chosen to reduce a certain Markov chain and then we will use the following notation. Let $\rho \in \mathcal{P}(\mathcal{S})$, then the clustered version that corresponds to the partitioning ρ will be denoted by adding a **tilde** and a **subscript** ρ such that the reduced Markov chain has state set $\tilde{\mathcal{S}}_\rho$, transition matrix $\tilde{P}_\rho(\pi)$ and distribution $\tilde{\pi}_\rho(\pi)$. If π is the stationary distribution then we drop the argument π and get transition matrix $\tilde{P}_\rho$ and distribution $\tilde{\pi}_\rho$.

Remark For two *disjunct* sets A and B we will use the notation $A + B$ to denote the unification instead of the more usual notation $A \cup B$. This is to emphasize the fact that the intersection $A \cap B$ is empty. (See for instance Expression (1.6) and Expression (1.13).

Example revisited

Next we try to apply the above clustering method to the example from the previous subsection shown in Figure 1.3. In the case of the example $\mathcal{S} = \{0, 1, 2\}$ and

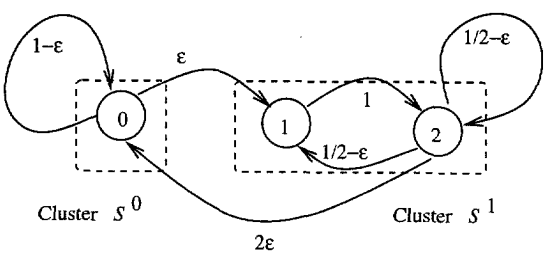

Figure 1.3: Graph of the Markov chain of the example.

$$P = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix}.$$

Suppose we have in mind to make one cluster from state 1 and 2 and one cluster from state 0, then we get the following disjoint partitioning $\rho \in \mathcal{P}(\mathcal{S})$:

$$\rho \ = \ (\mathcal{S}^0, \ \mathcal{S}^1) \ = \ (\{0\}, \ \{1, 2\}),$$

$$P = \begin{pmatrix} P^{(0,0)} & P^{(0,1)} \\ P^{(1,0)} & P^{(1,1)} \end{pmatrix},$$

where

$$\begin{aligned} P^{(0,0)} &= 1-\varepsilon\,, \\ P^{(0,1)} &= \begin{pmatrix} 0 & 2\varepsilon \end{pmatrix}, \\ P^{(1,0)} &= \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \\ P^{(1,1)} &= \begin{pmatrix} 0 & \frac{1}{2}-\varepsilon \\ 1 & \frac{1}{2}-\varepsilon \end{pmatrix}. \end{aligned}$$

Suppose we start again in the initial distribution $\pi(0)$:

$$\pi(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \pi^{(0)}(0) \\ \pi^{(1)}(0) \end{pmatrix} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right).$$

When we make the clustered version of this we get:

$$\tilde{\pi}_\rho(\pi(0)) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\tilde{p}_{\rho,00}(\pi(0)) = \frac{\| P^{(0,0)}\, \pi^{(0)}(0) \|_1}{\| \pi^{(0)}(0) \|_1} = \frac{(1-\varepsilon)\cdot 1}{1} = 1-\varepsilon\,,$$

$$\tilde{p}_{\rho,01}(\pi(0)) = \frac{\| P^{(0,1)}\, \pi^{(1)}(0) \|_1}{\| \pi^{(1)}(0) \|_1} = \frac{\begin{pmatrix} 0 & 2\varepsilon \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}}{\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \|_1} = \text{Not Determined}\,,$$

$$\tilde{p}_{\rho,10}(\pi(0)) = \frac{\| P^{(1,0)}\, \pi^{(0)}(0) \|_1}{\| \pi^{(0)}(0) \|_1} = \frac{\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \cdot 1 \|_1}{1} = \varepsilon\,,$$

$$\tilde{p}_{\rho,11}(\pi(0)) = \frac{\| P^{(1,1)}\, \pi^{(1)}(0) \|_1}{\| \pi^{(1)}(0) \|_1} = \frac{\| \begin{pmatrix} 0 & \frac{1}{2}-\varepsilon \\ 1 & \frac{1}{2}-\varepsilon \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \|_1}{\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \|_1} = \text{Not Determined}\,.$$

So we get:

$$\tilde{P}_\rho(\pi(0)) = \begin{pmatrix} 1-\varepsilon & \text{Not Determined} \\ \varepsilon & \text{Not Determined} \end{pmatrix}.$$

The problem that occurs here is that we have to divide 0 by 0, when that occurs it means that the distribution is such that the probability is 0 to be in one of the clusters. When this occurs "Not Determined" is put as an answer. This does not at all effect the generality and use of the clustering method, since at last we will always end up in a distribution where the probability to be in any cluster is and remains greater than 0.
So for the clustered Markov chain the first transition is as follows:

$$\tilde{\pi}_\rho(\pi(1)) = \begin{pmatrix} 1-\varepsilon & \text{Not Determined} \\ \varepsilon & \text{Not Determined} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1-\varepsilon \\ \varepsilon \end{pmatrix}.$$

Since we "know" $\pi(1)$ from the example in the previous subsection we can compute $\tilde{P}_\rho(\pi(1))$ in a similar way as we computed $\tilde{P}_\rho(\pi(0))$:

$$\begin{aligned}
\tilde{p}_{\rho,00}(\pi(1)) &= \frac{\| P^{(0,0)}\ \pi^{(0)}(1) \|_1}{\| \pi^{(0)}(1) \|_1} = \frac{(1-\varepsilon)\cdot(1-\varepsilon)}{1-\varepsilon} = 1-\varepsilon\ , \\
\tilde{p}_{\rho,01}(\pi(1)) &= \frac{\| P^{(0,1)}\ \pi^{(1)}(1) \|_1}{\| \pi^{(1)}(1) \|_1} = \frac{\begin{pmatrix} 0 & 2\varepsilon \end{pmatrix}\begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}}{\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \|_1} = 0\ , \\
\tilde{p}_{\rho,10}(\pi(1)) &= \frac{\| P^{(1,0)}\ \pi^{(0)}(1) \|_1}{\| \pi^{(0)}(1) \|_1} = \frac{\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \cdot (1-\varepsilon) \|_1}{1-\varepsilon} = \varepsilon\ , \\
\tilde{p}_{\rho,11}(\pi(1)) &= \frac{\| P^{(1,1)}\ \pi^{(1)}(1) \|_1}{\| \pi^{(1)}(1) \|_1} = \frac{\| \begin{pmatrix} 0 & \frac{1}{2}-\varepsilon \\ 1 & \frac{1}{2}-\varepsilon \end{pmatrix}\begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \|_1}{\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \|_1} = 1\ .
\end{aligned}$$

So we get:

$$\tilde{P}_\rho(\pi(1)) = \begin{pmatrix} 1-\varepsilon & 0 \\ \varepsilon & 1 \end{pmatrix}.$$

The problem of dividing 0 by 0 has vanished and will not return when we proceed in the same way. So for the clustered Markov chain the second transition is as follows:

$$\tilde{\pi}_\rho(\pi(2)) = \begin{pmatrix} 1-\varepsilon & 0 \\ \varepsilon & 1 \end{pmatrix} \begin{pmatrix} 1-\varepsilon \\ \varepsilon \end{pmatrix} = \begin{pmatrix} (1-\varepsilon)^2 \\ \varepsilon(1-\varepsilon)+\varepsilon \end{pmatrix}.$$

Now we can even proceed equivalently for the third transition: Since we "know" $\pi(2)$ from the example in the previous subsection we can compute $\tilde{P}_\rho(\pi(2))$ in a similar way as we

computed $\tilde{P}_\rho(\pi(1))$:

$$\begin{aligned}
\tilde{p}_{\rho,00}(\pi(2)) &= \frac{\| P^{(0,0)}\ \pi^{(0)}(2)\ \|_1}{\| \pi^{(0)}(2)\ \|_1} = \frac{(1-\varepsilon)\cdot(1-\varepsilon)^2}{(1-\varepsilon)^2} = 1-\varepsilon\ , \\
\tilde{p}_{\rho,01}(\pi(2)) &= \frac{\| P^{(0,1)}\ \pi^{(1)}(2)\ \|_1}{\| \pi^{(1)}(2)\ \|_1} = \frac{\begin{pmatrix} 0 & 2\varepsilon \end{pmatrix}\begin{pmatrix} \varepsilon(1-\varepsilon) \\ \varepsilon \end{pmatrix}}{\| \begin{pmatrix} \varepsilon(1-\varepsilon) \\ \varepsilon \end{pmatrix} \|_1} = \frac{2\varepsilon}{2-\varepsilon}\ , \\
\tilde{p}_{\rho,10}(\pi(2)) &= \frac{\| P^{(1,0)}\ \pi^{(0)}(2)\ \|_1}{\| \pi^{(0)}(2)\ \|_1} = \frac{\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}\cdot(1-\varepsilon)^2\ \|_1}{(1-\varepsilon)^2} = \varepsilon\ , \\
\tilde{p}_{\rho,11}(\pi(2)) &= \frac{\| P^{(1,1)}\ \pi^{(1)}(2)\ \|_1}{\| \pi^{(1)}(2)\ \|_1} = \frac{\| \begin{pmatrix} 0 & \frac{1}{2}-\varepsilon \\ 1 & \frac{1}{2}-\varepsilon \end{pmatrix}\begin{pmatrix} \varepsilon(1-\varepsilon) \\ \varepsilon \end{pmatrix} \|_1}{\| \begin{pmatrix} \varepsilon(1-\varepsilon) \\ \varepsilon \end{pmatrix} \|_1} = \frac{2-3\varepsilon}{2-\varepsilon}\ .
\end{aligned}$$

So we get:

$$\tilde{P}_\rho(\pi(2)) = \begin{pmatrix} 1-\varepsilon & \frac{2\varepsilon}{2-\varepsilon} \\ \varepsilon & \frac{2-3\varepsilon}{2-\varepsilon} \end{pmatrix}.$$

So for the clustered Markov chain the third transition is as follows:

$$\tilde{\pi}_\rho(\pi(3)) = \begin{pmatrix} 1-\varepsilon & \frac{2\varepsilon}{2-\varepsilon} \\ \varepsilon & \frac{2-3\varepsilon}{2-\varepsilon} \end{pmatrix}\begin{pmatrix} (1-\varepsilon)^2 \\ \varepsilon(1-\varepsilon)+\varepsilon \end{pmatrix} = \begin{pmatrix} (1-\varepsilon)^3+2\varepsilon^2 \\ \varepsilon(1-\varepsilon)^2+\varepsilon(2-\varepsilon) \end{pmatrix}.$$

This simple example shows us that when we have a non-varying constant Markov transition matrix P originally, we get varying Markov transition matrices $\tilde{P}_\rho(\pi(t))\ \ t=0,1,2,..$ if we compute the corresponding clustered variant. Originally we had the non-varying chain:

$$\begin{aligned}
\pi(1) &= P\ \pi(0)\ , \\
\pi(2) &= P\ \pi(1)\ , \\
\pi(3) &= P\ \pi(2)\ , \\
&\text{etcetera.}
\end{aligned} \tag{1.14}$$

But after clustering we "constructed" the following varying (non-homogeneous) chain:

$$\begin{aligned}
\tilde{\pi}_\rho(\pi(1)) &= \tilde{P}_\rho(\pi(0))\ \tilde{\pi}_\rho(\pi(0))\ , \\
\tilde{\pi}_\rho(\pi(2)) &= \tilde{P}_\rho(\pi(1))\ \tilde{\pi}_\rho(\pi(1))\ , \\
\tilde{\pi}_\rho(\pi(3)) &= \tilde{P}_\rho(\pi(2))\ \tilde{\pi}_\rho(\pi(2))\ , \\
&\text{etcetera.}
\end{aligned} \tag{1.15}$$

But, when the overall Markov chain reaches its stationary distribution π:

$$\pi = \frac{1}{7+2\varepsilon}\begin{pmatrix} 4 \\ 1+2\varepsilon \\ 2 \end{pmatrix},$$

the clustered version of the chain will also become stationary (homogeneous), since the clustered Markov transition matrix changes with the original distribution. Thus, when the original distribution becomes stationary the clustered Markov chain will become stationary also:

$$\begin{aligned}
\tilde{p}_{\rho,00}(\pi) &= \frac{\| P^{(0,0)}\,\pi^{(0)} \|_1}{\| \pi^{(0)} \|_1} = \frac{(1-\varepsilon)\cdot\frac{4}{7+2\varepsilon}}{\frac{4}{7+2\varepsilon}} = 1-\varepsilon\,, \\
\tilde{p}_{\rho,01}(\pi) &= \frac{\| P^{(0,1)}\,\pi^{(1)} \|_1}{\| \pi^{(1)} \|_1} = \frac{\begin{pmatrix} 0 & 2\varepsilon \end{pmatrix}\begin{pmatrix} \frac{1+2\varepsilon}{7+2\varepsilon} \\ \frac{2}{7+2\varepsilon} \end{pmatrix}}{\left\| \begin{pmatrix} \frac{1+2\varepsilon}{7+2\varepsilon} \\ \frac{2}{7+2\varepsilon} \end{pmatrix} \right\|_1} = \frac{4\varepsilon}{3+2\varepsilon}\,, \\
\tilde{p}_{\rho,10}(\pi) &= \frac{\| P^{(1,0)}\,\pi^{(0)} \|_1}{\| \pi^{(0)} \|_1} = \frac{\left\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}\cdot\frac{4}{7+2\varepsilon} \right\|_1}{\frac{4}{7+2\varepsilon}} = \varepsilon\,, \\
\tilde{p}_{\rho,11}(\pi) &= \frac{\| P^{(1,1)}\,\pi^{(1)} \|_1}{\| \pi^{(1)} \|_1} = \frac{\left\| \begin{pmatrix} 0 & \frac{1}{2}-\varepsilon \\ 1 & \frac{1}{2}-\varepsilon \end{pmatrix}\begin{pmatrix} \frac{1+2\varepsilon}{7+2\varepsilon} \\ \frac{2}{7+2\varepsilon} \end{pmatrix} \right\|_1}{\left\| \begin{pmatrix} \frac{1+2\varepsilon}{7+2\varepsilon} \\ \frac{2}{7+2\varepsilon} \end{pmatrix} \right\|_1} = \frac{3-2\varepsilon}{3+2\varepsilon}\,.
\end{aligned}$$

So we get the following stationary transition matrix for the clusters:

$$\lim_{t\to\infty} \tilde{P}_\rho(\pi(t)) = \tilde{P}_\rho = \begin{pmatrix} 1-\varepsilon & \frac{4\varepsilon}{3+2\varepsilon} \\ \varepsilon & \frac{3-2\varepsilon}{3+2\varepsilon} \end{pmatrix}. \tag{1.16}$$

The stationary distribution of the stationary clustered transition matrix is:

$$\tilde{\pi}_\rho = \frac{1}{7+2\varepsilon}\begin{pmatrix} 4 \\ 3+2\varepsilon \end{pmatrix}. \tag{1.17}$$

□

1.3.3 Long and Short Term Behaviour

Suppose again we have a Markov chain with a set of states $\mathcal{S}$ again divided in subsets:

$$\mathcal{S} = \mathcal{S}^0 + \mathcal{S}^1 + .. + \mathcal{S}^n, \tag{1.18}$$

but what is more, all subsets $\mathcal{S}^i$ correspond to a "mode" of behaviour. That means that whenever a cluster $\mathcal{S}^i$ is entered the probability of staying in the same cluster is very high, and the probability of a transition to another cluster is very low. So in fact each cluster can be seen as a "mode" of behaviour, in which the system will stay for a long time until a transition to another "mode" of behaviour will occur. For the transition matrix this implies that it has the following structure:

$$P_\varepsilon = \begin{pmatrix} P_\varepsilon^{(0,0)} & \varepsilon P_\varepsilon^{(0,1)} & . & . & \varepsilon P_\varepsilon^{(0,n)} \\ \varepsilon P_\varepsilon^{(1,0)} & P_\varepsilon^{(1,1)} & . & . & \varepsilon P_\varepsilon^{(1,n)} \\ . & . & & & . \\ . & . & & & . \\ \varepsilon P_\varepsilon^{(n,0)} & \varepsilon P_\varepsilon^{(n,1)} & . & . & P_\varepsilon^{(n,n)} \end{pmatrix}, \qquad (1.19)$$
$$\varepsilon \ll 1 .$$

In words: all off-diagonal sub-matrices have entries of an order ε smaller than the diagonal sub-matrices. For the sake of simplicity we will assume that the graph of non-zero transitions in P_0 consists of sub-graphs that are strongly connected themselves. This kind of Markov chain can be split into two kinds of behaviours. Each behaviour consists of one or more Markov chains on its own. There are $n+1$ sub-models each describing a short term behaviour in a mode. The short term behaviour of mode i consists of the set of states $\mathcal{S}^i$ and the corresponding transition matrix $P_0^{(i,i)}$. Next to the short term behaviours of the overall system, there also is one long term behaviour. The long term behaviour can be seen as the clustered version corresponding to the clusters of the several modes. The long term states are the separate modes. The long term behaviour describes how the overall system jumps between the several modes. The transition matrix of the clustered version is called the long term transition matrix. A special property of this kind of Markov chain is that the stationary distribution of the chain can be approximated easily: it needs much less computations to find a good approximation. In the sequel we will describe the approximation method. The approximation method splits the problem of finding the stationary distribution in more simple sub-problems: find the stationary distribution of the several modes and then find the stationary distribution of the long term behaviour.
Thus, the first step is to solve the stationary distributions π_0 for the behaviour in each short term mode, as if they are behaving independently:

$$\begin{aligned} \pi_0^{(i)} &= P_0^{(i,i)}\, \pi_0^{(i)} \quad \forall i , \\ \| \pi_0^{(i)} \|_1 &= 1 \quad \forall i , \end{aligned} \qquad (1.20)$$

and then use this stationary "distribution" (not really a distribution since $\| \pi_0 \|_1 = n+1$) to derive following approximated version of the clustered (=long term) stationary transition matrix:

$$\tilde{P}_\rho^{\text{appr}} = \tilde{P}_\rho(\pi_0) . \qquad (1.21)$$

Then solve with the approximated stationary version for the long term transition matrix its approximated stationary clustered distribution $\tilde{\pi}_{\rho}^{\text{appr}}$:

$$\tilde{\pi}_{\rho}^{\text{appr}} = \tilde{P}_{\rho}^{\text{appr}} \; \tilde{\pi}_{\rho}^{\text{appr}} \; . \tag{1.22}$$

Notice that in the current notation $\tilde{\pi}_{\rho,i}^{\text{appr}}$ are real valued scalars and $\pi_0^{(i)}$ are sub-vectors of the vector π_0. With this in mind, the approximated version of the overall stationary distribution can be formulated as:

$$\pi^{\text{appr}} = \begin{pmatrix} \tilde{\pi}_{\rho,0}^{\text{appr}} \pi_0^{(0)} \\ \tilde{\pi}_{\rho,1}^{\text{appr}} \pi_0^{(1)} \\ \cdot \\ \cdot \\ \tilde{\pi}_{\rho,n}^{\text{appr}} \pi_0^{(n)} \end{pmatrix} . \tag{1.23}$$

It will take some time from the start until the transient behaviour is over and the system reaches the stationary distribution. The time that it will take until the system is in its stationary behaviour can be estimated by computing the next to largest absolute value of the eigenvalue of the long term transition matrix. (Any stochastic matrix has an eigenvalue 1. Since we will only restrict ourselves to aperiodic Markov chains, there will be no other eigenvalues on the complex unit circle. A sufficient condition for a chain to be aperiodic is that all elements on the diagonal of the corresponding transition matrix are greater than zero. All other eigenvalues therefore lie completely within the unit circle. This follows from the Perron-Frobenius theorem. See [GS87] for further details on this subject.) So if the $n+1$ eigenvalues $\lambda_0, \lambda_1, .., \lambda_{n-1}, \lambda_n$, of the estimated long term transition matrix are such that:

$$|\lambda_0| \leq |\lambda_1| \leq .. \leq |\lambda_{n-1}| < \lambda_n = 1 \; . \tag{1.24}$$

Then a conservative estimation for the transient time $T_{\delta}^{\text{trans}}$ (in a number of time slots) until the distribution enters a δ-sphere around the stationary solution therefore is:

$$T_{\delta}^{\text{trans}} \leq \frac{\ln \delta - \ln n}{|\lambda_{n-1}| - 1} \; . \tag{1.25}$$

Example revisited

Next we try to apply the above approximation method to the example from the previous subsections again illustrated in Figure 1.4.
In the case of the example $\mathcal{S} = \{0, 1, 2\}$ and

$$P_{\varepsilon} = \begin{pmatrix} 1-\varepsilon & 0 & 2\varepsilon \\ \varepsilon & 0 & \frac{1}{2}-\varepsilon \\ 0 & 1 & \frac{1}{2}-\varepsilon \end{pmatrix} .$$

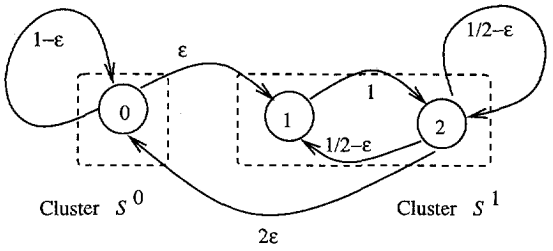

Figure 1.4: Graph of the Markov chain of the example.

Suppose we have in mind to make one cluster out of state 1 and 2, then we get the following disjoint partitioning of $\mathcal{S}$:

$$\begin{aligned} \mathcal{S} &= \{0\} \cup \{1,2\}\,, \\ P_\varepsilon &= \begin{pmatrix} P_\varepsilon^{(0,0)} & \varepsilon P_\varepsilon^{(0,1)} \\ \varepsilon P_\varepsilon^{(1,0)} & P_\varepsilon^{(1,1)} \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} P_\varepsilon^{(0,0)} &= 1-\varepsilon\,, \\ P_\varepsilon^{(0,1)} &= \begin{pmatrix} 0 & 2 \end{pmatrix}, \\ P_\varepsilon^{(1,0)} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ P_\varepsilon^{(1,1)} &= \begin{pmatrix} 0 & \frac{1}{2}-\varepsilon \\ 1 & \frac{1}{2}-\varepsilon \end{pmatrix}. \end{aligned}$$

So, as shown now, the example indeed has the "mode" structure. There are two short term behaviours or "modes", which are shown in Figure 1.5

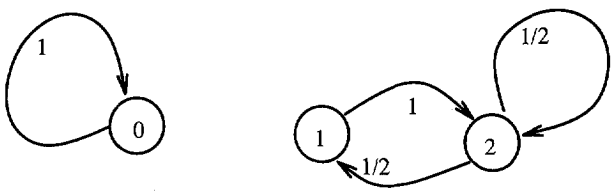

Figure 1.5: The two different modes: short term behaviour of the example.

Both modes correspond to the graphs of the following transition matrices:

$$\begin{aligned} P_0^{(0,0)} &= 1 , \\ P_0^{(1,1)} &= \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{pmatrix} . \end{aligned}$$

To approximate the stationary distribution we first determine the steady state distributions of both modes and solve:

$$\begin{aligned} \pi_0^{(0)} &= P_0^{(0,0)} \, \pi_0^{(0)} , \\ \pi_0^{(1)} &= P_0^{(1,1)} \, \pi_0^{(1)} , \\ \| \, \pi_0^{(i)} \, \|_1 &= 1 \quad \forall i. \end{aligned}$$

We solve and get:

$$\begin{aligned} \pi_0^{(0)} &= 1 , \\ \pi_0^{(1)} &= \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} . \end{aligned}$$

So we now have:

$$\pi_0 = \begin{pmatrix} 1 \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} .$$

Now we are going to make the approximation of the clustered stationary long term transition matrix:

$$\begin{aligned} \tilde{p}_{\rho,00}(\pi_0) &= \frac{\| \, P_\varepsilon^{(0,0)} \, \pi_0^{(0)} \, \|_1}{\| \, \pi_0^{(0)} \, \|_1} = \frac{(1-\varepsilon) \cdot 1}{1} = 1 - \varepsilon , \\ \tilde{p}_{\rho,01}(\pi_0) &= \frac{\| \, \varepsilon P_\varepsilon^{(0,1)} \, \pi_0^{(1)} \, \|_1}{\| \, \pi_0^{(1)} \, \|_1} = \frac{\begin{pmatrix} 0 & 2\varepsilon \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}}{\| \begin{pmatrix} 1 \\ 2 \end{pmatrix} \|_1} = \frac{4}{3}\varepsilon , \\ \tilde{p}_{\rho,10}(\pi_0) &= \frac{\| \, \varepsilon P_\varepsilon^{(1,0)} \, \pi_0^{(0)} \, \|_1}{\| \, \pi_0^{(0)} \, \|_1} = \frac{\| \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix} \cdot 1 \, \|_1}{1} = \varepsilon , \\ \tilde{p}_{\rho,11}(\pi_0) &= \frac{\| \, P_\varepsilon^{(1,1)} \, \pi_0^{(1)} \, \|_1}{\| \, \pi_0^{(1)} \, \|_1} = \frac{\| \begin{pmatrix} 0 & \frac{1}{2} - \varepsilon \\ 1 & \frac{1}{2} - \varepsilon \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \|_1}{\| \begin{pmatrix} 1 \\ 2 \end{pmatrix} \|_1} = 1 - \frac{4}{3}\varepsilon . \end{aligned}$$

So we get:

$$\tilde{P}_{\rho}^{\text{appr}} = \tilde{P}_{\rho}(\pi_0) = \begin{pmatrix} 1-\varepsilon & \frac{4}{3}\varepsilon \\ \varepsilon & 1-\frac{4}{3}\varepsilon \end{pmatrix}, \tag{1.26}$$

and indeed this is very close to the real stationary clustered transition matrix we derived earlier as shown in expression (1.16). The drawing in Figure 1.6 shows the approximation of the long term behaviour of the system of our example. We can now solve an approximation

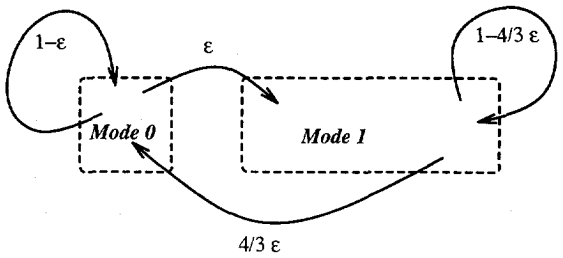

Figure 1.6: The long term behaviour of the example.

of the long term stationary distribution $\tilde{\pi}_{\rho}^{\text{appr}}$ from:

$$\tilde{\pi}_{\rho}^{\text{appr}} = \tilde{P}_{\rho}^{\text{appr}}\ \tilde{\pi}_{\rho}^{\text{appr}}\ ,$$

and get:

$$\tilde{\pi}_{\rho}^{\text{appr}} = \begin{pmatrix} \frac{4}{7} \\ \frac{1}{7} \\ \frac{2}{7} \end{pmatrix}. \tag{1.27}$$

When we compare this approximation with expression (1.17) then we can see the resemblance with the exact solution. From this approximation of the long term stationary distribution we can easily get the approximation of the overall distribution:

$$\pi^{\text{appr}} = \begin{pmatrix} \frac{4}{7} \cdot 1 \\ \frac{3}{7}\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \frac{4}{7} \\ \frac{1}{7} \\ \frac{2}{7} \end{pmatrix}. \tag{1.28}$$

This approximation should be compared with the exact overall distribution vector in expression (1.5).
In order to get an estimation for the transient time of the example we will estimate the time $T_{\frac{1}{100}}^{\text{trans}}$ (in number of time slots) until the distribution is in a $\frac{1}{100}$−neighbourhood of the stationary distribution. The eigenvalues of the long term transition matrix in expression (1.26) are:

$$\lambda_0 = 1 - \frac{7}{3}\varepsilon < 1 = \lambda_1.$$

So an estimation for the transient behaviour time is in this case:

$$T^{\text{trans}}_{\frac{1}{100}} \leq \frac{\ln 200}{\frac{7}{3}\varepsilon} = \frac{3 \ln 200}{7\varepsilon} \approx \frac{2.27}{\varepsilon}. \tag{1.29}$$

□

1.3.4 Conclusions

In this subsection we described certain aspects of homogeneous discrete-time finite Markov chains. In models of real life systems the number of states in a Markov chain is large. As the number of states is large in general, reduction and simplification of the models is a must in order to keep the amount of computations limited even if a large computer is available. A tool to simplify a large Markov chain is to divide the states into clusters. The construction of simplified Markov chains from bigger ones by clustering states was explained. Another way to simplify a Markov chain even more is to divide the states into clusters such that all clusters correspond to different modes of behaviour. If the last division of states is possible, the overall stationary distribution can be approximated by a simple method that needs less computations than finding the exact stationary distribution. Of course, in such a case, the approximation found will not be equal to the exact stationary distribution. However, for purposes in real life models in which the number of states is very high the profit of finding the exact solution is low in comparison with the extra costs of the effort of finding the exact solution. For these models the approximation method in this section can be considered as very good tool to find a stationary distribution. The division of Markov chains into clusters or even into long and short term behaviours can be very useful for the simplification of large Markov chains.

Chapter 2

Serial Production Lines with Product Failures Only.

In the next three chapters we will examine serial production lines. In order to be clear we will first describe how a serial production line is defined.

A *serial production line* is a production line that has exactly one input buffer and one output buffer and in which each buffer and each machine have precisely one input and one output arc, except for the input and output buffer. As a consequence a *serial production line* consists of m machines and $m+1$ buffers which can all be given a number in correspondence with their consecutive order in the line: $\mathcal{M} = \{M_1, M_2, ..., M_m\}$ and $\mathcal{B} = \{B_1, B_2, ..., B_{m+1}\}$. A picture of a serial production line is shown in Figure 2.1.

A serial production line can have several features. We divide serial production lines in several classes depending on the kind of failures that can occur inside the line. In Chapter 1 we have seen that we will make a difference between product failures and machine failures. Therefore we can distinguish production lines with *product failures only*, production lines with *machine failures only* and production lines with both *product and machine failures*. In the following chapters we will deal with these different classes and show how approximations of the steady state distribution of these classes of production lines can be

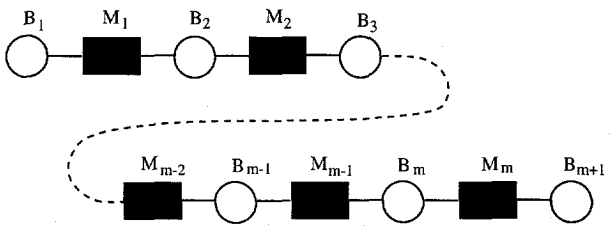

Figure 2.1: A serial production line.

calculated. The central concept used for approximation of the steady state distributions is decomposition of a serial production line into smaller serial production lines consisting of two machines only. The decomposition method for production lines with product failures only and the method for the production lines with machine failures only look similar but are totally different. The two different methods will be described and explained in this chapter, Chapter 2, and in the next chapter, Chapter 3. After that a combination of the two methods will be described in Chapter 4.

2.1 Description of a Serial Production line with Product Failures Only

The serial manufacturing systems in this section are defined by the following eight assumptions:

1. One machine *cycle* consists of three activities:
   - If present, pick up a product part from the buffer in front of the machine. In case no part is present the machine is called *starved.*
   - If a product part was picked up, manipulate the product part. Manipulation completes with a probability referred to as *completion probability.* Manipulation fails with a probability referred to as *failure probability.*
   - If there is space available and the manipulation is successfully completed put the product part into the next buffer, otherwise put the product part back into the last position of the buffer in front of the machine. (This procedure is because we assume that there is not a position to store a product part inside a machine, this position is considered to be the last position *in* the previous buffer.) In case there is no room in the next buffer the machine is called *blocked.*
2. The time is slotted with a fixed slot duration T_{slot}. Within a single time slot the machines perform a single *cycle* in parallel. However, because of conflict situations that can occur in case buffers are empty or full, we have to specify an order in which machines take and put product parts in buffers. Therefore we specify the following order: *In a conflict situation at buffer B_i the next machine M_i always has priority over the previous machine M_{i-1}.*
3. At the end of a time slot there are no product parts within any machine. Therefore it is impossible for the product parts to accumulate within machines.
4. The capacity of each buffer B_i is $N_i \in I\!N^+$, $i = 1, .., m$. We will assume that the first and the last buffers have infinite capacity: $N_1 = N_{m+1} = \infty$.
5. Machine M_1 is never starved, machine M_m is never blocked.

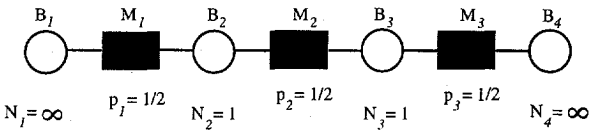

Figure 2.2: The simple three-machine production line.

6. Machine M_i has *completion probability* p_i, and *failure probability* $q_i = 1-p_i$ $i = 1, .., m$.

7. The failures at the machines are independent with respect to time, independent with respect to each other and independent with respect to the present or past state of the production line.

Remark. Note that by these assumptions the *completion probability* p_i of a machine M_i is equal to the average production rate of machine M_i if the machine is never starved or blocked. Note also that by means of these assumptions, in fact we obtain a model of a serial queueing network with blocking that has machines with discrete *geometrically distributed service times.*

2.2 A Simple Example

We have shown the most simple three-machine production line in Figure 2.2. The parameters for the machines and buffers in the system are also mentioned in the figure. We can see in the picture that we are dealing with three machines $\mathcal{M} = \{M_1, M_2, M_3\}$ and four buffers $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$. The capacities of buffers B_1, B_2, B_3, B_4 are ∞, 1, 1, ∞ respectively. In the figure we also mention the parameters p_1, p_2 and p_3 corresponding to the features of the service times of machines M_1, M_2 and M_3 respectively. These parameters will be called the *completion probabilities* of their machines.
As mentioned in the previous chapter we will assume that the simple production line can be modelled by means of a *homogeneous finite discrete time Markov Chain.* First of all this implies that we assume that time is divided in time slots and that transitions take place at the end of each slot.
The procedure for this three-machine production line *inside one time slot* consists of the three following *cycles*:

M_3-cycle At the beginning of each time slot machine M_3 starts its cycle. That means that it looks in upstream buffer B_3 if there is a product part waiting. If there is no product part waiting the cycle of machine M_3 ends. If a product part is waiting then it takes the part and the cycle continues by manipulating the product part. Manipulation succeeds with *completion probability* $p_3 = \frac{1}{2}$ and fails with *failure probability* $q_3 = 1 - p_3 = \frac{1}{2}$. If manipulation fails then machine M_3 puts the product part back in upstream buffer B_3 and

stops its cycle. If manipulation succeeds then machine M_3 looks in the downstream buffer B_4 if there is room available for the manipulated product part. By assumption the last buffer is never full. Therefore there is always room available in buffer B_4. Since buffer B_4 is not full machine M_3 can put the product part in downstream buffer B_4 and ends its cycle.

M_2-cycle Next it is machine M_2's turn to do one cycle. It looks in upstream buffer B_2 if there is a product part waiting. If there is no product part waiting the cycle of machine M_2 ends. If a product part is waiting then it takes the part and the cycle continues by manipulating the product part. Manipulation succeeds with *completion probability* $p_2 = \frac{1}{2}$ and fails with *failure probability* $q_2 = 1 - p_2 = \frac{1}{2}$. If manipulation fails then machine M_2 puts the product part back in upstream buffer B_2 and stops its cycle. If manipulation succeeds then machine M_2 looks in the downstream buffer B_3 if there is room available for the manipulated product part. If buffer B_3 is full, which is the case here if there is already $N_3 = 1$ product part in B_3, then machine M_2 puts the product part back in upstream buffer B_2 and stops its cycle. If buffer B_2 is not full then machine M_2 puts the product part in downstream buffer B_3 and ends its cycle as well.

M_1-cycle Finally it is machine M_1's turn to do one cycle. It looks in upstream buffer B_1 if there is a product part waiting. By assumption there are always product parts waiting in the first buffer B_1 and therefore machine M_1 takes a product part and continues its cycle by manipulating it. Manipulation succeeds with *completion probability* $p_1 = \frac{1}{2}$ and fails with *failure probability* $q_1 = 1 - p_1 = \frac{1}{2}$. If manipulation fails then machine M_1 puts the product part back in upstream buffer B_1 and stops its cycle. If manipulation succeeds then machine M_1 looks in the downstream buffer B_2 if there is room available for the manipulated product part. If buffer B_2 is full, which is the case here if there is already $N_2 = 1$ product part in B_2, machine M_1 puts the product part back in upstream buffer B_1 and stops its cycle. If buffer B_2 is not full then machine M_1 puts the product part in downstream buffer B_2 and ends its cycle as well.

At this moment each machine has completed one cycle and the time slot ends. A new time slot can start and the whole procedure starts all over again.

□

2.3 Problem Statement

The following problems, concerning the analysis and synthesis of serial production lines with product failures only, will be dealt with in this chapter:

1. Given the completion probabilities p_i, $i = 1, .., m$ and the buffer capacities N_i, $i = 2, .., m-1$, find a method to approximate the stationary distribution of the system.

From the approximation of the stationary distribution we can find an approximation for the *average production rate* which is the number that production line managers are most interested in. This problem is referred to as the problem of analysis.

2. Given a required and beforehand specified average production rate R of the system, find the optimal completion probabilities $p_1, .., p_m$ such that the *total workforce* $W = \sum_i p_i$ is minimized. A system that satisfies this condition is called *dynamically balanced.*

3. Find the optimal distribution of the total buffer capacity $C = \sum_i N_i$, C given, over the separate buffers B_i, $i = 1, .., m-1$, under the condition that the resulting system will be dynamically balanced.

In this chapter a non-asymptotic theory is developed that approximates the solutions to problems 1-3. Asymptotic versions of the theory in the sense that it examines limiting behaviour of the systems described here can be found in [MT90]. In the article of this last reference a method is derived for similar production lines in which the machine failure probabilities q_i are very small. Exact analytic solutions for the production lines that we describe here have, as far as the author knows, never been obtained. Problem 1 is referred to as the problem of analysis. Problems 2 and 3 are referred to as the problem of synthesis. In the next sections we deal with the problems of analysis and synthesis respectively.

2.4 Analysis

2.4.1 First Approach to the Problem

The situation at the beginning of time slot $[nT_{\text{slot}}, (n+1)T_{\text{slot}})$ is fixed with the contents of each buffer. Since the system has $m+1$ buffers of which buffers B_1 and B_{m+1} have an infinite capacity and are assumed to contain an infinite number of product parts all the time, the following state space X will be assigned to the system:

$$X = \{0, 1, .., N_2\} \times \{0, 1, .., N_3\} \times ..... \times \{0, 1, .., N_m\}. \tag{2.1}$$

Now we can describe the system in terms of the stochastic sequence: $\{x_0,\ x_1,\ x_2, ...\}$ where $x_i \in X$ for $i = 0, 1, 2, ...$. When we examine this stochastic sequence, we conclude that it satisfies the following Markov-property (for the theory of Markov chains see [GS87]):

$$\mathbb{P}\text{r}(x_n = u | x_0, x_1, .., x_{n-1}) = \mathbb{P}\text{r}(x_n = u | x_{n-1}),\ \forall\ u \in X,\ \forall\ n \in I\!N. \tag{2.2}$$

Further we notice that the chain X is homogeneous:

$$\mathbb{P}\text{r}(x_{n+1} = u | x_n = v) = \mathbb{P}\text{r}(x_1 = u | x_0 = v),\ \forall\ u, v \in X,\ \forall\ n \in I\!N. \tag{2.3}$$

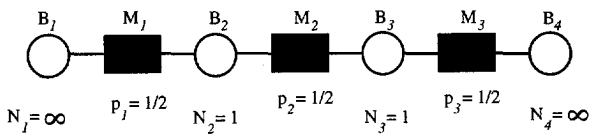

Figure 2.3: The simple three-machine production line.

Now we define:

$$Y = \{0, 1, 2, .., A-1\}, \tag{2.4}$$

where

$$A = (N_2 + 1).(N_3 + 1).(N_4 + 1)....(N_m + 1).$$

State space X and set Y have the same finite number of elements. Therefore there exists a bijection F from X to Y. There exist as many bijective mappings F as there are permutations of A elements. Although many bijective mappings F exist, any of these mappings will do for our purposes. By this means we create a new state space Y of which the elements are natural numbers. The stochastic sequence $\{x_0,\ x_1,\ x_2, ...\}$ has the image under F: $\{F(x_0),\ F(x_1),\ F(x_2), ...\} = \{y_0,\ y_1,\ y_2, ...\}$. The sequence $\{y_0,\ y_1,\ y_2, ...\}$ describes the evolution of the system as well.

We will show that for our problem that there exists a transition matrix $P = (p_{i,j})$ with the transition probabilities:

$$p_{i,j} = \mathbb{P}\mathrm{r}(y_1 = i | y_0 = j),\ i, j\ \in\ Y. \tag{2.5}$$

From this transition matrix we can compute the stationary distribution or steady state π of the chain by solving the equations:

$$\begin{aligned} P\,\pi &= \pi, \\ \sum_{i \in Y} \pi_i &= 1, \\ \pi_i &\geq 0, \quad \forall\, i\ \in\ Y. \end{aligned} \tag{2.6}$$

2.4.2 The Simple Example Revisited

Consider again the simple example from a previous section, Section 2.2, as again shown in Figure 2.3. In this simple example we have four buffers $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ which have the capacities $\infty, 1, 1, \infty$ respectively and three machines $\mathcal{M} = \{M_1, M_2, M_3\}$ which have the completion probabilities $\frac{1}{2},\ \frac{1}{2},\ \frac{1}{2}$ respectively.
What determines the state at the beginning of each time slot, is only the contents of buffers B_2 and B_3. The contents of buffer B_2 and B_3 is either 0 or 1 product parts. This means for our example that the state set X is equal to: $X = \{0, 1\} \times \{0, 1\}$. Since it is inconvenient

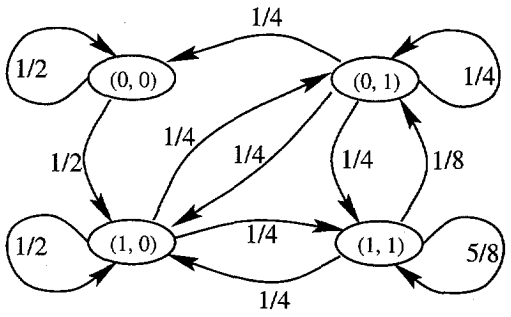

Figure 2.4: The Markov chain corresponding to the simple three-machine line.

to deal with pairs of numbers we make a new state set Y with four natural numbers only: $Y = \{0, 1, 2, 3\}$. Next we define a bijection F that maps X on Y as follows:

$$\begin{array}{rcl} F(\ (0,0)\ ) & = & 0, \\ F(\ (0,1)\ ) & = & 1, \\ F(\ (1,0)\ ) & = & 2, \\ F(\ (1,1)\ ) & = & 3. \end{array} \tag{2.7}$$

Then we can setup the transition matrix $P = (p_{i,j})$ by determining all the transition probabilities $p_{i,j} = \mathbb{P}\mathrm{r}(y_{n+1} = i | y_n = j)$:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{5}{8} \end{pmatrix}. \tag{2.8}$$

The corresponding Markov chain is shown in Figure 2.4.
Next we solve $P\pi = \pi$ and get the stationary distribution solution:

$$\pi = \frac{1}{11} \begin{pmatrix} 1 \\ 2 \\ 4 \\ 4 \end{pmatrix}. \tag{2.9}$$

From the stationary distribution we can easily derive the average production rate of the simple three-machine production line. We look at the output of the production line and we see that there can only be output just after states 1 and 3. This is because these are the only two states with a product part in buffer B_3. If at the beginning of a time slot we are in either states 1 or 3 we have the probability of $p_3 = \frac{1}{2}$ that there will be a product part added to the output buffer B_4 at the end of the time slot. This means

that in the case of our example we have an average production rate R *exactly* equal to $p_3(\pi_1 + \pi_3) = \frac{1}{2}(\frac{2}{11} + \frac{4}{11}) = \frac{3}{11}$ product part per time slot.

□

In principle this example shows a method to find the steady state of every possible production line of this nature. However, the example deals with a *simple* production line with four states only. Notice that the number of states increases *rapidly* with the buffer capacities. For example if we have m buffers (apart from input and output buffers) with capacity N each then we can count $(N+1)^m$ states in total (see state space X in Expression (2.1)). For a production line with five buffers of capacity 100 each, which is not an uncommon situation, we have to deal with more than 10^{10} states in the state set. An amount of 10^{10} states in the state set will give some problems for the straightforward method that we have just described:

1. We have to *determine* each element of a $10^{10} \times 10^{10}$ transition matrix P that corresponds to the production line. (Even if most of the elements are 0.)

2. We have to *store* a $10^{10} \times 10^{10}$ transition matrix P.

3. We have to *compute* a stationary distribution vector π of length 10^{10} corresponding to the $10^{10} \times 10^{10}$ transition matrix P.

Another drawback of the straightforward method is that it only computes the stationary distribution, but does not give insight into more simple relationships between variables and does not give insight in the sensitivity of the average production rate for changes in the parameters.

We conclude that because of the large dimensions and the complexity of the transition matrix P it is a problem to find an analytic solution for π in general. In order to obtain more insight into the problem we will examine the smallest possible system first: the two-machine system.

2.4.3 The Two-Machine System

We now examine the simple case of two machines and three buffers of which there is an input buffer and an output buffer that have infinite capacity. In this case we have two machines M_1 and M_2 with completion probabilities p_1 and p_2 respectively. Since there is only one main buffer with finite capacity we will simply call it B. Its capacity is assumed to be equal to N. A picture is shown in Figure 2.5.
The state space Y of this simple system is:

$$Y = \{0, 1, 2, 3, .., N-1, N\}. \tag{2.10}$$

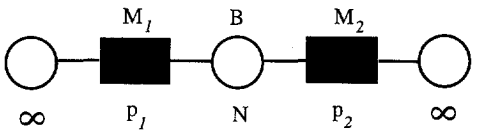

Figure 2.5: A two-machine production line.

The transition matrix of the system has therefore the dimensions $(N+1) \times (N+1)$. The transition matrix has the following nice tri-diagonal structure, which is common for similar queueing models:

$$P = \begin{pmatrix} q_1 & q_1p_2 & 0 & 0 & \cdot & \cdot & 0 \\ p_1 & p_1p_2+q_1q_2 & q_1p_2 & 0 & & & \cdot \\ 0 & p_1q_2 & p_1p_2+q_1q_2 & q_1p_2 & & & \cdot \\ \cdot & & \ddots & \ddots & \ddots & & \cdot \\ \cdot & & & \ddots & \ddots & \ddots & \cdot \\ \cdot & & & & p_1q_2 & p_1p_2+q_1q_2 & q_1p_2 \\ 0 & \cdot & \cdot & \cdot & 0 & p_1q_2 & 1-q_1p_2 \end{pmatrix}. \tag{2.11}$$

Of course we are interested in the stationary distribution of this system. Therefore we solve the right eigenvector π corresponding to eigenvalue 1 of transition matrix P, as defined in Equation (2.6). The author succeeded to find an analytic formula for the stationary distribution of this system:

$$\begin{aligned} \pi_0 &= q_1 . \frac{1-\alpha}{1-\frac{q_1}{q_2}.\alpha^{N+1}}, \\ \pi_i &= \frac{\pi_0}{q_2}.\alpha^i, \quad \text{for } i = 1,..,N, \end{aligned} \tag{2.12}$$

where

$$\alpha = \frac{p_1q_2}{p_2q_1}. \tag{2.13}$$

Notice that the stationary distribution in (2.12) is not defined if p_1 equals p_2. If $p_1 = p_2 = p$ (and consequently $q_1 = q_2 = q$), then the stationary distribution is:

$$\begin{aligned} \pi_0 &= \frac{q}{N+q}, \\ \pi_i &= \frac{1}{N+q}, \quad \text{for } i = 1,..,N. \end{aligned} \tag{2.14}$$

For the sake of abbreviation we introduce the functions Ψ_N for $N \in \mathbb{N}^+$:

$$\Psi_N : [0,1] \times [0,1] \mapsto [0,1],$$

$$\Psi_N(x,y) \stackrel{\text{def}}{=} \begin{cases} \dfrac{1-\alpha}{1-\frac{y}{x}\alpha^{N+1}}, & \text{for } (x,y) \in (0,1)^2 \text{ with } x \neq y\ , \\ \dfrac{1}{N+x}, & \text{for } (x,y) \in [0,1]^2 \text{ with } x = y\ , \\ \max(2-N,0).(1-x), & \text{for } (x,y) \in [0,1]^2 \text{ with } \dfrac{y(1-x)}{y-x} = 0\ , \\ 1, & \text{for } (x,y) \in [0,1]^2 \text{ with } \dfrac{x(1-y)}{x-y} = 0\ , \end{cases} \tag{2.15}$$

where

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

The exact properties of the functions Ψ_N are described in Appendix A in Section A.1 of this thesis. The functions Ψ_N are continuous except for the points (0,0) and (1,1). This will soon be made clear.

With (2.15) we can reformulate Equations (2.12), (2.13) and (2.14) as follows:

$$\begin{aligned} \pi_0 &= q_1.\Psi_N(q_2,q_1), \\ \pi_i &= \left(\frac{p_2 q_1}{p_1 q_2}\right)^{N-i}.\Psi_N(q_1,q_2), \quad \text{for } i = 1,..,N. \end{aligned} \tag{2.16}$$

The interpretation of the function Ψ_N follows from π_N in Expression (2.16): Ψ_N represents the probability that buffer B with capacity N is full if the corresponding two-machine line behaves stationary. In order to get more feeling for the functions Ψ_N we have done some numerical computations and made some three-dimensional graphs of the first four of them. The functions Ψ_1, Ψ_2, Ψ_3 and Ψ_4 are shown in Figures 2.6, 2.7, 2.8 and 2.9 respectively.

If we examine formula (2.16) and see the interpretation of Ψ_N in terms of π_0 or π_N then the exceptions on the edge of the functions Ψ_N will become clear:

- In case $q_1 = 1$ or $q_2 = 0$ the contents of the buffer can only decrease, which means that all states except the empty one are transient. Therefore $\pi_0 = 1$ and $\pi_N = 0$ holds.

- In case $q_1 = 0$ or $q_2 = 1$ the contents of the buffer can only increase, which means that all states except the full one are transient. Therefore $\pi_0 = 0$ and $\pi_N = 1$ holds.

- Very special cases are the case $q_1 = 0$ and $q_2 = 0$ and the case $q_1 = 1$ and $q_2 = 1$. In these two cases the contents of the buffer will remain unchanged in every time cycle. In these cases it is impossible to derive the stationary distribution of the system from q_1 and q_2 only! We have to know the initial contents of the buffer to derive a "steady state". In this section we assume the initial contents of the buffer to be 1. Therefore $\pi_0 = 0$ and $\pi_N = 0$ (if $N > 1$, $\pi_N = 1$ if $N = 1$) in these two cases.
- The discontinuity of Ψ_N for $N > 1$ in (0,0) is clear since if $q_1 = \delta$ and $q_2 = 0$ then $\pi_N = 0$ and if $q_1 = 0$ and $q_2 = \delta$ then $\pi_N = 1$. $(0 < \delta \ll 1)$.
- The discontinuity of Ψ_N in (1,1) is clear since if $q_1 = 1 - \delta$ and $q_2 = 1$ then $\pi_N = 1$ and if $q_1 = 1$ and $q_2 = 1 - \delta$ then $\pi_N = 0$. $(0 < \delta \ll 1)$.

With (2.16) we can derive the average production rate R of the two machines one buffer system. We can derive the average production rate by two means: from the input and from the output of the system.

First we derive it by means of the input. Therefore we examine the occurrence that machine M_1 succeeds in taking, manipulating, and putting a product part in the buffer B within one time slot. Taking a product part will not be a problem, since there are product parts available at any time. Manipulating a product part within one time slot happens with a completion probability p_1. Putting a product part in buffer B can only happen if there is room available in it, this is always the case, except when buffer B was full at the beginning of the operation, and machine M_2 fails manipulating its taken product part. Hence we can formulate the following expression for the average input production rate:

$$R^{\text{input}} = p_1.(1 - q_2.\pi_N) = p_1.(1 - q_2.\Psi(q_1, q_2)). \tag{2.17}$$

Next we derive the average production rate by means of the output. Therefore we examine the occurrence that machine M_2 succeeds in taking a product part out of buffer B, manipulating it and putting it out of the system. Taking a product part out of buffer B is only possible if the buffer is not empty at the beginning of the time slot. Manipulating the product part within one time slot happens with a completion probability p_2. Delivering the product part to outside the system will not be a problem, because we presume infinite storage room at the output of the system. Hence we can formulate the next expression for the average output production rate:

$$R^{\text{output}} = p_2.(1 - \pi_0) = p_2.(1 - q_1.\Psi(q_2, q_1)). \tag{2.18}$$

Of course it should be true that the average input rate R^{input} in (2.17) equals R^{output} in (2.18). This is obvious if we realize the necessity of the property that says: "what comes in must come out". Indeed we can prove that the following always holds:

$$\begin{aligned} R^{\text{input}} &= R^{\text{output}} \\ p_1.(1 - q_2.\Psi(q_1, q_2)) &= p_2.(1 - q_1.\Psi(q_2, q_1)). \end{aligned} \tag{2.19}$$

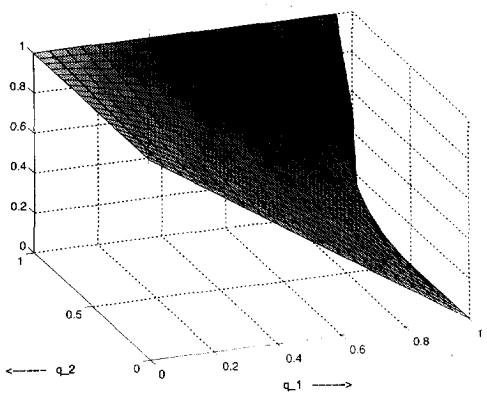

Figure 2.6: Three-Dimensional plot of Ψ_1.

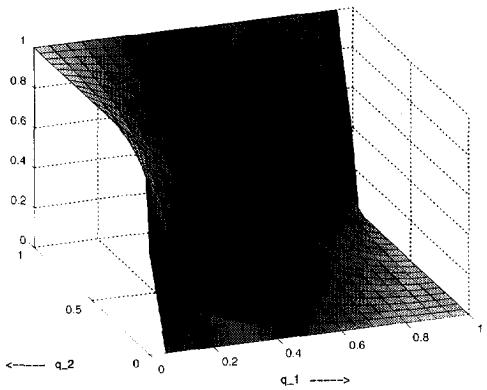

Figure 2.7: Three-Dimensional plot of Ψ_2.

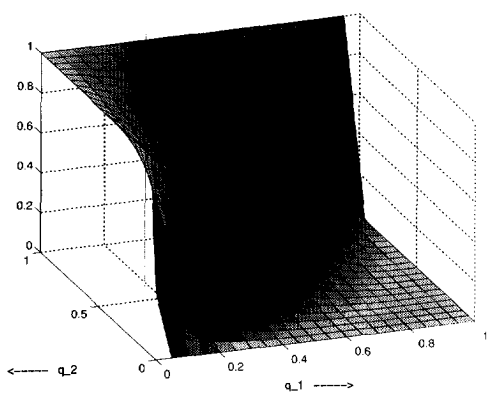

Figure 2.8: Three-Dimensional plot of Ψ_3.

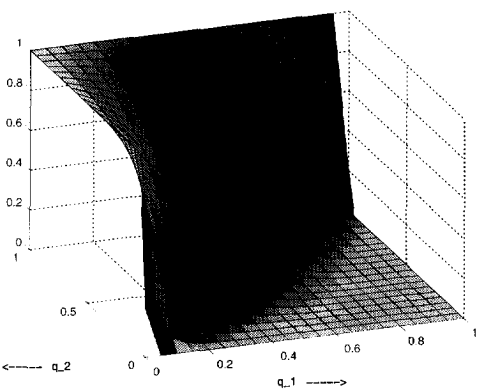

Figure 2.9: Three-Dimensional plot of Ψ_4.

The production rate $R = R^{\text{input}} = R^{\text{output}}$ as a function of p_1 and p_2 will play an important role in further analysis. Therefore we will define the functions R_N for $N \in \mathbb{N}^+$ as follows:

$$R_N : [0,1] \times [0,1] \mapsto [0,1],$$

$$\begin{array}{rcl} R_N(x,y) & \stackrel{\text{def}}{=} & x.(1-(1-y).\Psi_N(1-x,1-y)) \\ & = & y.(1-(1-x).\Psi_N(1-y,1-x)) \end{array} \begin{array}{l} = \\ = R_N(y,x) \end{array} \tag{2.20}$$

The functions R_N and its most important properties can be found in Appendix A in Section A.2. A technical demonstration of the symmetry of R_N with respect to its arguments is shown under "R_N property 4" in Theorem 9. In order to obtain some feeling we have done some numerical computations and made three-dimensional representations of the first four functions R_N. We can see the functions R_1, R_2, R_3 and R_4 in Figures 2.10, 2.11, 2.12 and 2.13 respectively. We can see clearly from these figures that our functions R_N all have a "pyramid shape" of which the "edge" sharpens as N increases.

The blockage and starvation probabilities are easy to find: machine M_1 is blocked when B is full and machine M_2 fails to complete a product part. Machine M_2 is starved when buffer B is empty.

$$\begin{array}{rcl} \Pr(M_1 \text{ is blocked}) & = & q_2.\pi_N \\ \Pr(M_2 \text{ is starved}) & = & \pi_0 \end{array} \begin{array}{l} = q_2.\Psi_N(q_1,q_2), \\ = q_1.\Psi_N(q_2,q_1). \end{array} \tag{2.21}$$

The results derived in the preceding paragraphs can be extended with the following three observations:

1. From the point of view of the average steady state production rate, the two machines one buffer system is equivalent to a single *aggregated* machine characterized by:

$$p_{\text{aggregation}} = p_2.(1-q_1\Psi_N(q_2,q_1)) = p_1.(1-q_2\Psi_N(q_1,q_2)). \tag{2.22}$$

2. If $N = 1$ (which is the smallest possible buffer capacity), the aggregated machine is characterized by:

$$p_{\text{aggregation}} = \frac{p_1 p_2}{1-q_1 q_2}, \quad \text{if } N = 1. \tag{2.23}$$

(See also "R_N property 3" in Theorem 8, from Appendix A in Section A.2.)
This is easy to verify since in this case we are dealing with the following transition matrix:

$$P = \begin{pmatrix} q_1 & q_1 p_2 \\ p_1 & 1-q_1 p_2 \end{pmatrix}. \tag{2.24}$$

3. If $N \to \infty$ (which is the "largest" possible buffer capacity), the aggregated machine is characterized by:

$$p_{\text{aggregation}} = \lim_{N\to\infty} R_N(p_1,p_2) = \min(p_1,p_2). \tag{2.25}$$

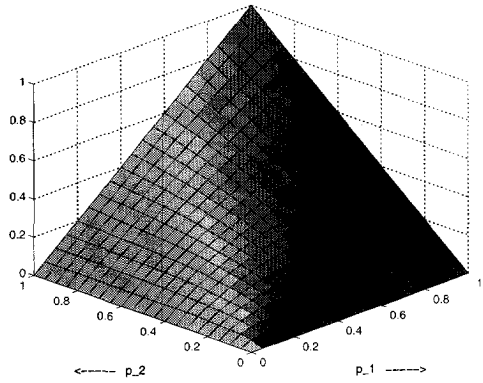

Figure 2.10: Three-Dimensional plot of R_1.

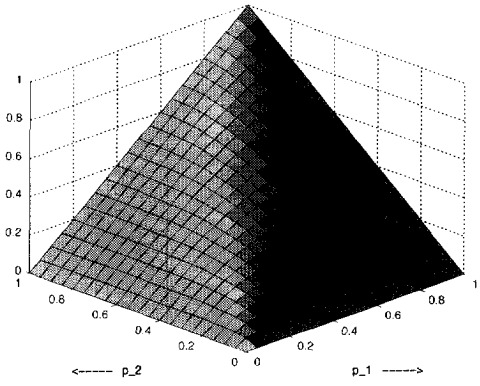

Figure 2.11: Three-Dimensional plot of R_2.

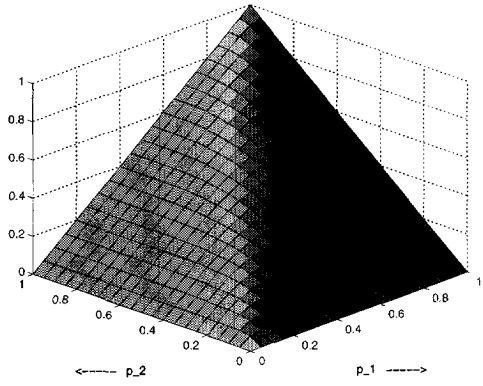

Figure 2.12: Three-Dimensional plot of R_3.

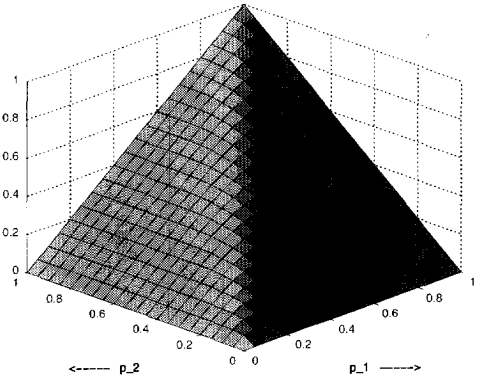

Figure 2.13: Three-Dimensional plot of R_4.

(See "R_N property 10" in Theorem 15 from Appendix A in Section A.2 for an extensive proof.)
This was to be expected since it is obvious that the machine with the smallest completion probability determines the total production rate when there is no limit to the buffer capacity.

2.4.4 Simplification of the General Model

After this thorough investigation of the two-machine production line we return to our more general m-machine production line. Consider again a line with m machines and $m+1$ buffers. Our aim is to derive an algorithm that approximates the stationary distribution of the total production line, using the results of the two-machine production lines from the previous subsection.

Consider again the state space X corresponding to the m-machine production line mentioned in Expression (2.1):

$$X = \{0, 1, .., N_2\} \times \{0, 1, .., N_3\} \times ..... \times \{0, 1, .., N_m\}. \tag{2.26}$$

This state space X corresponds via a bijective mapping F with state a space Y that consists of natural numbers, see Expression (2.4):

$$Y = \{0, 1, 2, .., A-1\}. \tag{2.27}$$

In order to "reduce" the Markov chain corresponding to the m-machine production line we make *partitionings* of the state space X. (Note that for reduction of Markov chains we will use the notation with a *tilde* and the corresponding partitioning as a *subscript* "$\tilde{(.)}_\rho$", introduced in Subsection 1.3.2). For the reduction we first we define the following subsets:

$$X_j^i \stackrel{\text{def}}{=} \{(x_2, .., x_i, .., x_m) \in X \mid x_i = j\}, \qquad i \in \{2, 3, .., m\},\ j \in \{0, 1, .., N_i\}, \tag{2.28}$$

by means of which we can define *partitionings* τ^i as follows:

$$\begin{array}{lcl} \tau^i & \in & \mathcal{P}(X), \\ \tau^i & \stackrel{\text{def}}{=} & (X_0^i, X_1^i, .., X_{N_i}^i). \end{array} \tag{2.29}$$

(See also the definition in (1.13) for notation.) The partitionings τ^i of state space X induce by means of bijective mapping F corresponding partitionings ρ^i of the state set Y. If

$$Y_j^i \stackrel{\text{def}}{=} F(\, X_j^i\,), \qquad i \in \{2, 3, .., m\},\ j \in \{0, 1, .., N_i\}, \tag{2.30}$$

then similarly

$$\begin{array}{lcl} \rho^i & \in & \mathcal{P}(Y), \\ \rho^i & \stackrel{\text{def}}{=} & (Y_0^i, Y_1^i, .., Y_{N_i}^i). \end{array} \tag{2.31}$$

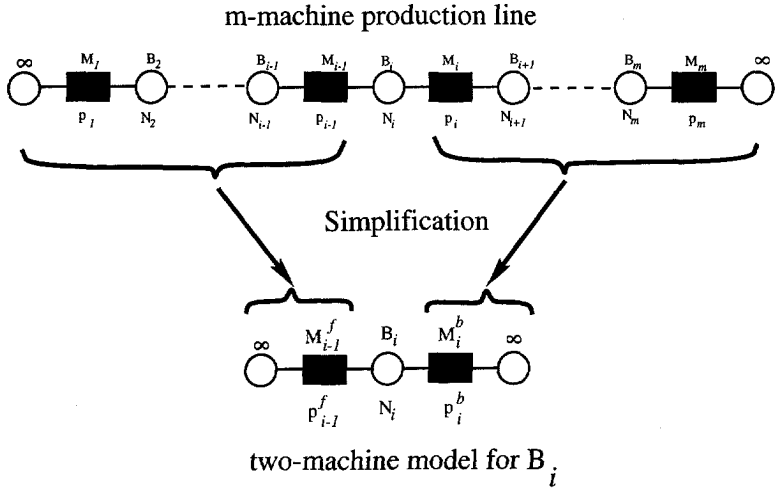

Figure 2.14: The simplified model for buffer B_i with virtual machines M_{i-1}^f and M_i^b.

We denote the overall Markov transition matrix corresponding to state set Y by P. Let us now focus on the transition matrix of the reduced Markov chain $\tilde{P}_{\rho^i}$. The structure of this transition matrix $\tilde{P}_{\rho^i}$ is *tri-diagonal* as well. This is obvious since it is impossible for the contents of a buffer to increase or decrease by more than one part within one time slot:

$$\tilde{P}_{\rho^i} = \begin{pmatrix} \tilde{p}_{\rho^i,00} & \tilde{p}_{\rho^i,01} & & & & \\ \tilde{p}_{\rho^i,10} & \tilde{p}_{\rho^i,11} & \tilde{p}_{\rho^i,12} & & & \\ & \tilde{p}_{\rho^i,21} & \tilde{p}_{\rho^i,22} & \tilde{p}_{\rho^i,23} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \tilde{p}_{\rho^i,N_i,N_i-1} & \tilde{p}_{\rho^i,N_i,N_i} \end{pmatrix}. \tag{2.32}$$

Now we compare the transition matrix $\tilde{P}_{\rho^i}$ of Expression (2.32) with the transition matrix of the two-machine production line in Expression (2.11). The two matrices both have a tri-diagonal structure. We ask ourselves if there exist values for the parameters p_1 and p_2 that make P in (2.11) exactly equal to $\tilde{P}_{\rho^i}$. Some analysis shows however that in general such values for p_1 and p_2 can *not* be found since in general:

$$\begin{array}{lcll} \tilde{p}_{\rho^i,n-1,n} & \neq & \tilde{p}_{\rho^i,n,n+1}, & \text{for } n \in \{1,2,..,N_i-1\}, \\ \tilde{p}_{\rho^i,n,n-1} & \neq & \tilde{p}_{\rho^i,n+1,n}, & \text{for } n \in \{2,..,N_i-1\}. \end{array} \tag{2.33}$$

This is in contrast to all transition matrices $P = (p_{i,j})$ of Expression (2.11), since there it

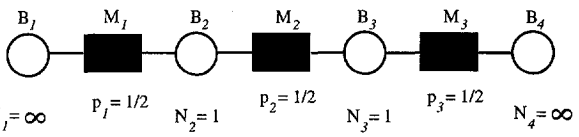

Figure 2.15: The simple three-machine production line.

is clear that equality holds indeed:

$$\begin{array}{rclcll} p_{n-1,n} & = & p_{n,n+1} & = & q_1 p_2, & \text{for all } n \in \{1,2,..,N-1\}, \\ p_{n,n-1} & = & p_{n+1,n} & = & p_1 q_2, & \text{for all } n \in \{2,..,N-1\}. \end{array} \tag{2.34}$$

The *simplification* of our m-machine production line model lies exactly in the difference between the two transition matrices since *instead of* the exact transition matrix $\tilde{P}_{\rho^i}$ we take the "best fitting" transition matrix P of the two-machine production line model from Expression (2.11) for certain "best values" of parameters p_1 and p_2. These "best values" of parameters p_1 and p_2 in order to approximate $\tilde{P}_{\rho^i}$ will be denoted by p_{i-1}^f and p_i^b respectively. How these "best values" are computed will be explained in the following subsection. Both p_{i-1}^f and p_i^b will have the interpretation of completion probabilities of *virtual aggregated machines* M_{i-1}^f and M_i^b respectively. A picture with the interpretation of the virtual machines is shown in Figure 2.14.

2.4.5 The Simple Example Revisited

Consider again the simple example from a previous section, Section 2.2, as again shown in Figure 2.15. In this simple example we have four buffers $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ which have the capacities $\infty, 1, 1, \infty$ respectively and three machines $\mathcal{M} = \{M_1, M_2, M_3\}$ which have the completion probabilities $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$ respectively.
For this system we found that the state set is $X = \{0,1\} \times \{0,1\}$. We also determined a set of natural numbers $Y = \{0,1,2,3\}$, and a bijective mapping $F : X \mapsto Y$ as follows:

$$\begin{array}{rcl} F(\ (0,0)\ ) & = & 0, \\ F(\ (0,1)\ ) & = & 1, \\ F(\ (1,0)\ ) & = & 2, \\ F(\ (1,1)\ ) & = & 3. \end{array} \tag{2.35}$$

For this simple system we derived already the exact transition matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{5}{8} \end{pmatrix}, \tag{2.36}$$

and the exact stationary distribution

$$\pi = \frac{1}{11}\begin{pmatrix} 1 \\ 2 \\ 4 \\ 4 \end{pmatrix}. \tag{2.37}$$

In the previous subsection we have defined the clusters X_j^i as follows

$$X_j^i \stackrel{\text{def}}{=} \{(x_2, .., x_i, .., x_m) \in X \mid x_i = j\}, \qquad i \in \{2, 3, .., m\},\ j \in \{0, 1, .., N_i\},$$

For our example this makes:

$$\begin{array}{rclcrcl} X_0^2 & = & \{(0,0),\ (0,1)\} & , & X_1^2 & = & \{(1,0),\ (1,1)\}, \\ X_0^3 & = & \{(0,0),\ (1,0)\} & , & X_1^3 & = & \{(0,1),\ (1,1)\}, \end{array}$$

which makes the following subsets of Y:

$$\begin{array}{rclcrcl} Y_0^2 & = & \{0,\ 1\} & , & Y_1^2 & = & \{2,\ 3\}, \\ Y_0^3 & = & \{0,\ 2\} & , & Y_1^3 & = & \{1,\ 3\}. \end{array}$$

The definition of *partitionings* τ^i is as follows:

$$\begin{array}{rcl} \tau^i & \in & \mathcal{P}(X), \\ \tau^i & \stackrel{\text{def}}{=} & (X_0^i, X_1^i, .., X_{N_i}^i). \end{array}$$

This makes for our example that

$$\begin{array}{rcl} \tau^2 & = & (\{(0,0),\ (0,1)\},\ \{(1,0),\ (1,1)\}\ ), \\ \tau^3 & = & (\{(0,0),\ (1,0)\},\ \{(0,1),\ (1,1)\}\ ), \end{array}$$

and therefore with mapping F we get the partitionings

$$\begin{array}{rcl} \rho^2 & = & (\{0,\ 1\},\ \{2,\ 3\}\ ), \\ \rho^3 & = & (\{0,\ 2\},\ \{1,\ 3\}\ ). \end{array}$$

Having determined the partitionings of the state sets X and Y we can derive the reduced Markov chains corresponding to them:

$$\tilde{P}_{\rho^2} = \begin{pmatrix} \frac{1}{2} & \frac{3}{16} \\ \frac{1}{2} & \frac{13}{16} \end{pmatrix}, \tag{2.38}$$

and

$$\tilde{P}_{\rho^3} = \begin{pmatrix} \frac{3}{5} & \frac{1}{3} \\ \frac{2}{5} & \frac{2}{3} \end{pmatrix}. \tag{2.39}$$

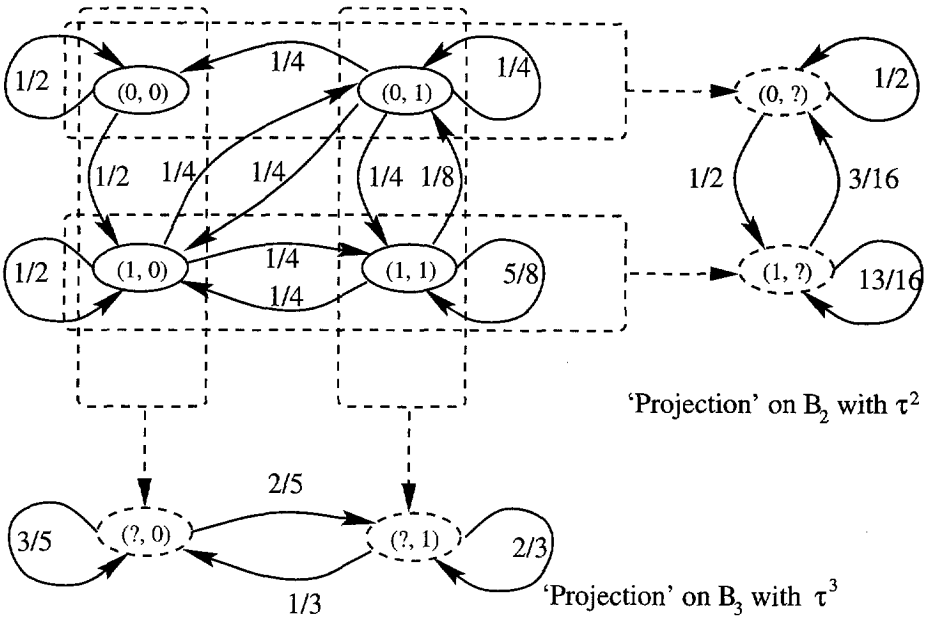

Figure 2.16: The Markov chain with partitionings τ^2 and τ^3 for "projections" on buffer B_2 and B_3.

By the nature of the partitionings we can consider the last two reduced Markov chains with transition matrices $\tilde{P}_{\rho^2}$ and $\tilde{P}_{\rho^3}$ in (2.38) and (2.39) as a kind of *projections* of the original Markov chain with transition matrix P. To illustrate this we have made a picture of the two "projections" in Figure 2.16.

It is clear that the "projected Markov chains" with transition matrices $\tilde{P}_{\rho^2}$ and $\tilde{P}_{\rho^3}$ both are again Markov chains corresponding to a single queue with capacity 1. Unfortunately it is *not* true that the "projected Markov chains" both correspond to a two-machine production line model with a buffer of capacity 1. If the "projected Markov chains" both did correspond to a two-machine production line model with a buffer of capacity 1 then we could freely decompose the three-machine production line into two two-machine production lines as shown in Figure 2.17. That means that we could find parameters p_2^f, q_2^f with $p_2^f + q_2^f = 1$, and p_2^b, q_2^b with $p_2^b + q_2^b = 1$ such that

$$\tilde{P}_{\rho^2} = \begin{pmatrix} \frac{1}{2} & \frac{3}{16} \\ \frac{1}{2} & \frac{13}{16} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \cdot p_2^b \\ \frac{1}{2} & 1 - \frac{1}{2} \cdot p_2^b \end{pmatrix},$$

$$\tilde{P}_{\rho^3} = \begin{pmatrix} \frac{3}{5} & \frac{1}{3} \\ \frac{2}{5} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} q_2^f & \frac{1}{2} \cdot q_2^f \\ p_2^f & 1 - \frac{1}{2} \cdot q_2^f \end{pmatrix}.$$

The first of these two equations has a solution for $p_2^b = \frac{3}{8}$ but the second equation leads to a contradiction since both $\frac{3}{5} = q_2^f$ and $\frac{1}{3} = \frac{1}{2} \cdot q_2^f$ cannot hold simultaneously. It turns out that we cannot put the "projected Markov chains" into a two-machine production line model without making small adaptations.

□

2.4.6 A Closer Look at the Simplified Model

It is our aim to derive the average production rate R of the total production line. Since we are dealing with a serial production line the average production rate R of the total production line equals the average production rate of any machine within the line.

When the average production rate of the total line equals $\frac{1}{2}$ product part per time slot, then the mean production rate of each machine within the line equals $\frac{1}{2}$ product part per time slot as well. If two adjacent machines would have different average production rates, then there would be a continuous increase or decrease of product parts in a buffer between two machines. A continuous increase or decrease of product parts in a buffer is impossible, because the buffers have limited capacities. Therefore the average production rates of the separate machines have to be equal to each other and equal to the average production rate of the total production line. This principle is denoted by the term *conservation of flow.*

Now we look at the average production rate of machine M_i. Machine M_i produces a

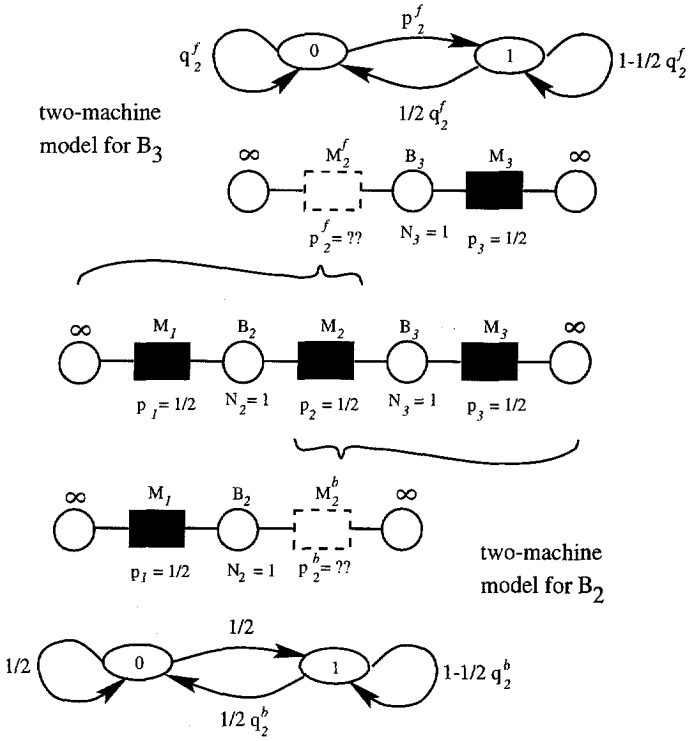

Figure 2.17: The decomposition of the system into two two-machine production lines.

product part within one time slot if the following three situations occur simultaneously: machine M_i is not starved, completes manipulation of the product part and is not blocked. Therefore, if we *assume* that the processes of blocking and starvation for the same machine are independent, the expression for the average production rate of machine M_i is as follows:

$$R = \mathbb{P}\mathrm{r}(M_i \text{ is not starved}).p_i.\mathbb{P}\mathrm{r}(M_i \text{ is not blocked}). \tag{2.40}$$

We define the average input production rate p_i^b, and the average output production rate p_i^f of machine M_i as follows:

$$\begin{array}{rcl} p_i^b & \stackrel{\text{def}}{=} & p_i.\mathbb{P}\mathrm{r}(M_i \text{ is not blocked}), \\ p_i^f & \stackrel{\text{def}}{=} & p_i.\mathbb{P}\mathrm{r}(M_i \text{ is not starved}). \end{array} \tag{2.41}$$

These definitions will soon be elucidated.
From (2.40) and (2.41) follows directly:

$$R = \frac{p_i^f p_i^b}{p_i}. \tag{2.42}$$

Now we look at the total production line as two machines M_{i-1}^f,M_i^b and one buffer B_i in between, as explained in the Subsection 2.4.4 and shown in Figure 2.14. An important question is: what should we take for the completion probabilities that belong to the aggregated machines M_{i-1}^f and M_i^b? The reasoning that we follow in order to determine these completion probabilities is as follows:

The completion probability p_{i-1}^f of the imaginary machine M_{i-1}^f should be equal to the average production rate of machine M_{i-1} (which is equal to R) with blocking due to buffer B_i left out of consideration. Similarly, the completion probability p_i^b of the imaginary machine M_i^b is equal to the average production rate of machine M_i (which is also equal to R) with the starvation due to buffer B_i left out of consideration.

The term "left out of consideration" here will be translated into a conditional probability. This conditional probability transforms into a division under the *assumption* that the event of a successfull production of a part and phenomena such as blocking or starvation are independent. Therefore:

$$\begin{array}{rcl} p_{i-1}^f & = & \dfrac{R}{\mathbb{P}\mathrm{r}(M_{i-1} \text{ is not blocked})}, \\ p_i^b & = & \dfrac{R}{\mathbb{P}\mathrm{r}(M_i \text{ is not starved})}. \end{array} \tag{2.43}$$

When we examine definition (2.41), we notice that it is in accordance with Expressions (2.40) and (2.43).

For the sake of abbreviation we define the complements of p_i^f and p_i^b as follows:

$$\begin{array}{rcl} q_i^f & \stackrel{\text{def}}{=} & 1 - p_i^f, \\ q_i^b & \stackrel{\text{def}}{=} & 1 - p_i^b. \end{array} \tag{2.44}$$

Now we have all ingredients to apply Expressions (2.20) and (2.21) to our imaginary two machine one buffer system. We get the following from Expression (2.20):

$$R = R_{N_i}(p_{i-1}^f, p_i^b) = R_{N_i}(p_i^b, p_{i-1}^f), \quad \text{for } i \in \{2, .., m\}. \tag{2.45}$$

From Expression (2.21) we get:

$$\begin{array}{rcl} \mathbb{P}\text{r}(M_{i-1} \text{ is blocked}) & = & q_i^b.\Psi_{N_i}(q_{i-1}^f, q_i^b), \quad \text{for } i \in \{2, .., m\}, \\ \mathbb{P}\text{r}(M_m \text{ is blocked}) & = & 0 \end{array} \tag{2.46}$$

and:

$$\begin{array}{rcl} \mathbb{P}\text{r}(M_i \text{ is starved}) & = & q_{i-1}^f.\Psi_{N_i}(q_i^b, q_{i-1}^f), \quad \text{for } i \in \{2, .., m\}, \\ \mathbb{P}\text{r}(M_1 \text{ is starved}) & = & 0 \end{array} \tag{2.47}$$

Expressions (2.40) up to (2.47) form the basic expressions of this chapter.
When we now combine Expression (2.46) and Expression (2.47) with $i = i - 1$ in (2.41) we get:

$$\begin{array}{rcll} p_i^f & = & p_i.(1 - q_{i-1}^f.\Psi_{N_i}(q_i^b, q_{i-1}^f)), & \text{for } i \in \{2, .., m\}, \\ p_i^b & = & p_i.(1 - q_{i+1}^b.\Psi_{N_{i+1}}(q_i^f, q_{i+1}^b)), & \text{for } i \in \{1, .., m-1\}, \\ p_1^f & = & p_1, \qquad p_m^b = p_m. & \end{array} \tag{2.48}$$

For the sake of abbreviation we define the following functions $\Omega_{c,N}$ for $c \in (0, 1]$ and $N \in \mathbb{N}^+$:

$$\begin{array}{rcl} \Omega_{c,N} & : & [0,1] \times [0,1] \mapsto [0,c], \\ \Omega_{c,N}(x,y) & \stackrel{\text{def}}{=} & c.(1 - y.\Psi_N(x,y)). \end{array} \tag{2.49}$$

The exact properties of these functions are mentioned in Appendix A in Section A.3. By means of these functions $\Omega_{c,N}$ we can describe (2.48) as follows:

$$\begin{array}{rcll} p_i^f & = & \Omega_{p_i,N_i}(q_i^b, q_{i-1}^f), & \text{for } i \in \{2, .., m\}, \\ p_i^b & = & \Omega_{p_i,N_{i+1}}(q_i^f, q_{i+1}^b), & \text{for } i \in \{1, .., m-1\}, \\ p_1^f & = & p_1, \qquad p_m^b = p_m. & \end{array} \tag{2.50}$$

The question is now how to compute p_i^f and p_i^b for all $i \in \{1, .., m\}$ in such a way that they satisfy Expression (2.50)? Do such real numbers p_i^f and p_i^b exist anyway? If a solution exists, will this solution be unique? Does a solution of Expression (2.50) automatically satisfy Expression (2.45)?

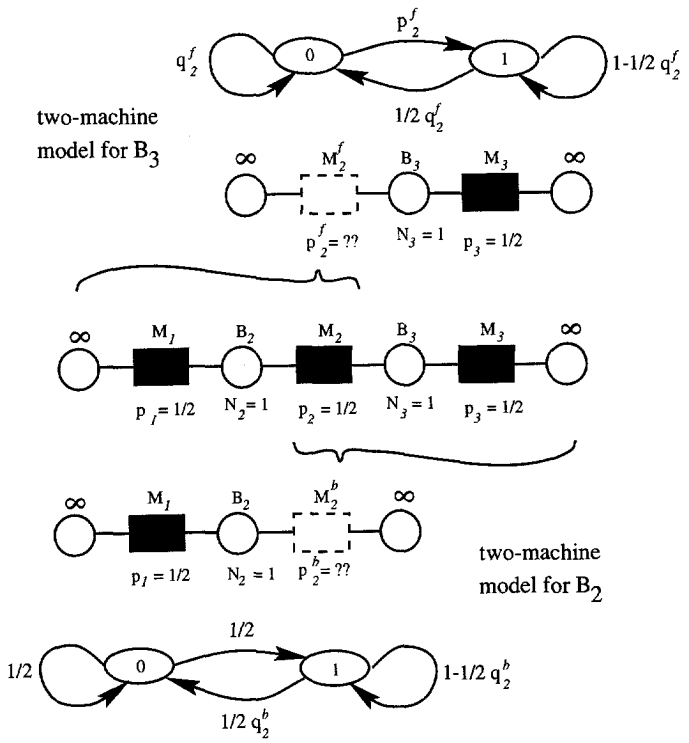

Figure 2.18: The decomposition of the system into two two-machine production lines.

2.4.7 The Simple Example Revisited

Consider again the simple example from a previous section, Section 2.2. In this simple example we have four buffers $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ which have the capacities $\infty, 1, 1, \infty$ respectively and three machines $\mathcal{M} = \{M_1, M_2, M_3\}$ which have the completion probabilities $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$ respectively. A picture of the example, this time together with its decomposition, is shown in Figure 2.18. For this system we have made a simplified model of which the decomposition is again shown in Figure 2.18. Because of the conservation of flow property we want that both two-machine production lines have exactly the same average production rate. Therefore we get:

$$\begin{aligned} R_{N_2}(p_1, p_2^b) &= R, \\ R_{N_3}(p_2^f, p_3) &= R. \end{aligned}$$

Next to these two equations we require that according to Expression (2.42), which links the two two-machine models, we get:

$$R = \frac{p_2^f \cdot p_2^b}{p_2}.$$

Since $N_2 = N_3 = 1$ we can use for our equations that

$$R_1(x, y) = \frac{xy}{x + y - xy}.$$

(See also "R_N property 3" in Theorem 8, from Appendix A in Section A.2.)
If we use this and substitute the parameters $p_1 = p_2 = p_3 = \frac{1}{2}$ then we get the following equations:

$$\begin{aligned}
\frac{\frac{1}{2}p_2^b}{\frac{1}{2} + p_2^b - \frac{1}{2}p_2^b} &= R, \\
\frac{\frac{1}{2}p_2^f}{\frac{1}{2} + p_2^f - \frac{1}{2}p_2^f} &= R, \\
\frac{p_2^f \cdot p_2^b}{\frac{1}{2}} &= R.
\end{aligned}$$

It is not difficult to solve this system of equations and to show that there exists a *unique* solution $(p_2^b, p_2^f) \in [0, 1]^2$:

$$\begin{aligned}
p_2^b &= p_2^f = \frac{\sqrt{3} - 1}{2}, \\
R &= 2 - \sqrt{3}.
\end{aligned}$$

What we have found now is a "best decomposition" of a three-machine production line into two two-machine production lines. The "exact" solution of the three-machine production line can now be compared with the approximation of the two two-machine models.
First of all with the two two-machine model approximation we obtain an average production rate $R = 2 - \sqrt{3}$. Earlier we have computed the exact average production rate $R = \frac{3}{11}$ of the three-machine production line. We conclude that the relative error we make in our example by decomposition for the average production rate equals $\frac{\frac{3}{11} - (2 - \sqrt{3})}{\frac{3}{11}} \cdot 100\% \approx 1.75\%$.

□

2.4.8 Solving for p_i^f and p_i^b in General

In total we have $2m$ equations in Expression (2.50). We have m unknowns of the form p_i^f and we have m unknowns of the form p_i^b, which makes $2m$ unknowns in total. The conclusion is that we have a system of $2m$ equations with $2m$ unknowns. A routine numerical procedure might be used to solve the system. Because of the particular structure of the equations the following special procedure is developed.

In the procedure we start with an array $p_i^f(0)$ of conservative estimates of p_i^f. We suggest to start with the most conservative estimates $p_i^f(s) = 0$ for $i = 1, 2, .., m$. After this initialization we begin with a cycle that makes new arrays of estimates $p_i^b(s)$ and $p_i^f(s+1)$ from the old array of estimates $p_i^f(s)$. The cycle is divided into two stages: first the backward aggregation and then the forward aggregation.

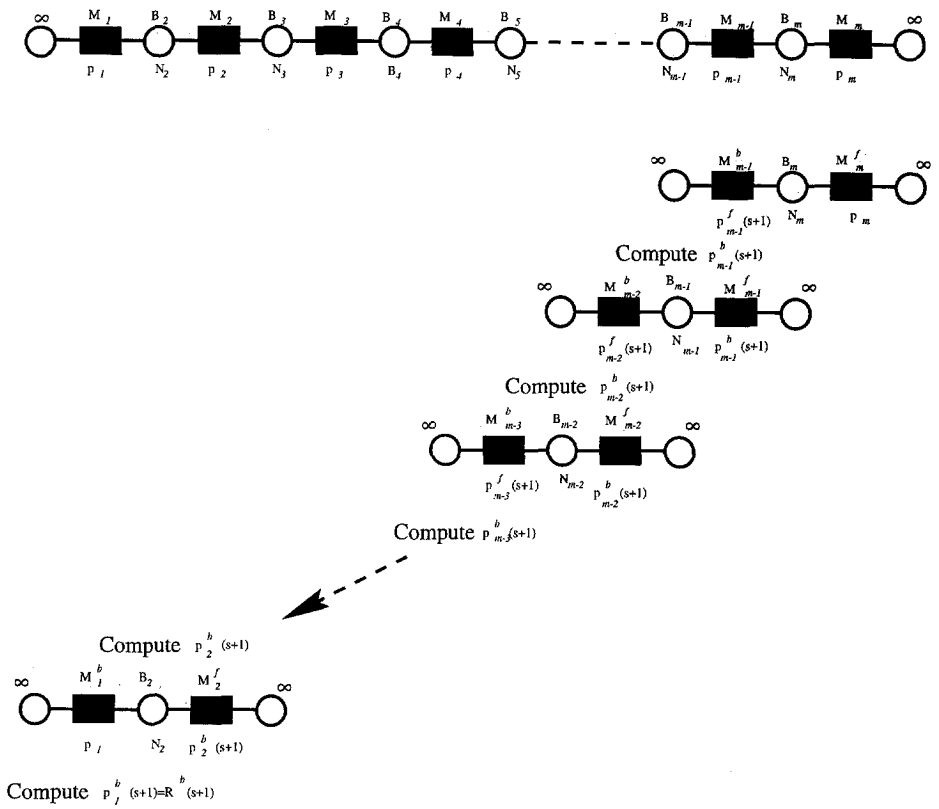

Figure 2.19: Illustration of the backward aggregation procedure.

The backward aggregation procedure

The *backward aggregation* procedure is illustrated in Figure 2.19. It makes an array of new estimates $p_i^b(s)$, $i = m-1, m-2, .., 1$ using the information in array $p_i^f(s)$, $i \in \{1, ..., m\}$. All new estimates in the backward aggregation process are obtained by applying the second equation of (2.50) repeatedly. First the backward aggregation makes a new estimate $p_{m-1}^b(s)$ by "aggregating" the last two machines M_m^b $(= M_m)$ and M_{m-1}. Next it makes a new estimate p_{m-2}^b by "aggregating" machines M_{m-1}^b and M_{m-2}. This process of aggregation continues until it ends with "aggregating" machines M_2^b and M_1 to find $p_1^b(s)$, which is the backward estimate $R^b(s+1)$ of the overall production rate R of the total production line.

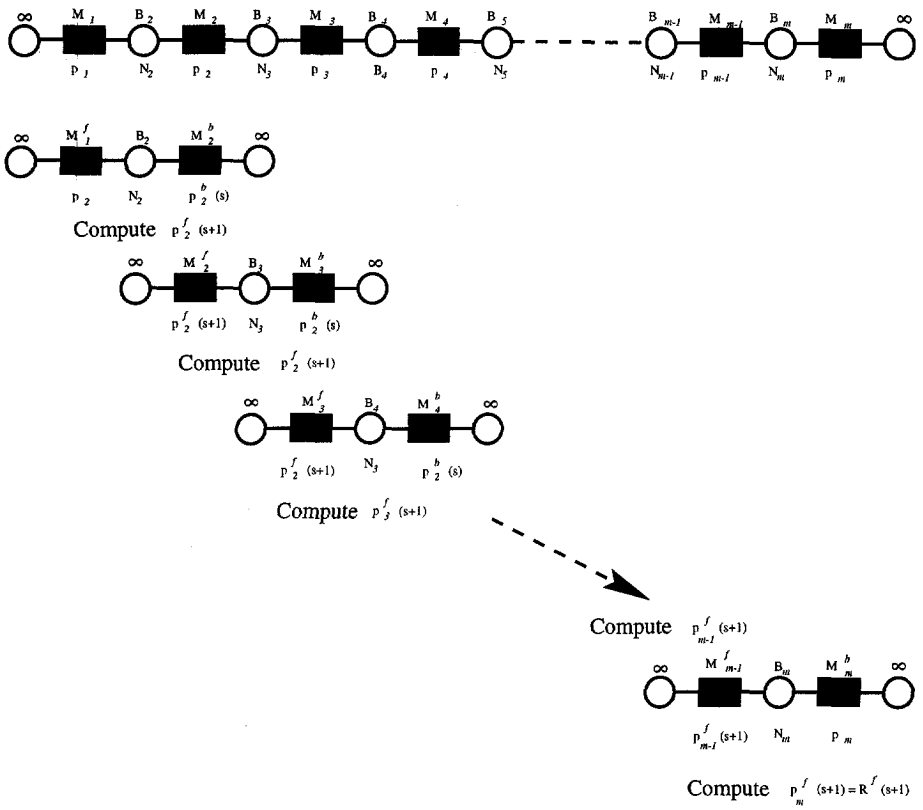

Figure 2.20: Illustration of the forward aggregation procedure.

The forward aggregation procedure

The *forward aggregation* procedure is illustrated in Figure 2.20. It makes an array of new estimates $p_i^f(s+1),\ i = 2, 3, .., m$ using information in array $p_i^b(s),\ i \in \{1, ..., m\}$ (which we have made in the previous backward aggregation process). All new estimates in the forward aggregation process are obtained by applying the first equation of (2.50) repeatedly. First the forward aggregation process makes a new estimate $p_2^f(s+1)$ by "aggregating" the first two machines $M_1^f\ (= M_1)$ and M_2. Next it makes a new estimate $p_3^f(s+1)$ by "aggregating" machines M_2^f and M_3. This process of aggregation continues until it ends with "aggregating" machines M_{m-1}^f and M_m to find estimate $p_m^f(s+1)$, which is the forward estimate $R^f(s+1)$ of the overall production rate R of the total production line.

To resume we formulate the algorithm as follows:

$$\textbf{Boundary conditions}:$$

$$p_1^f(s) = p_1, \quad p_m^b(s) = p_m, \quad \forall\, s \in I\!N,$$

$$\textbf{Initial conditions}:$$

$$p_i^f(0) = 0, \quad \forall\, i \in \{2, .., m\}\,, \tag{2.51}$$

$$\textbf{Backward aggregation for } i = m-1, m-2, .., 1:$$

$$p_i^b(s) = \Omega_{p_i, N_{i+1}}(q_i^f(s), q_{i+1}^b(s)),$$

$$\textbf{Forward aggregation for } i = 2, 3, .., m:$$

$$p^f(s+1) = \Omega_{p_i, N_i}(q_i^b(s), q_{i-1}^f(s+1)).$$

Illustrations of the forward and backward aggregation can be found in Figures 2.20 and 2.19.

We have proved that all variables $p_i^f(s)$ and $p_i^b(s)$ in this algorithm converge to a limit as s increases. The fact that the limit is a solution to Equations (2.50) has been proved as well. As a consequence it has been shown that the solution to Equations (2.50) does exist. We have also proved that this solution to Equations (2.50) is unique and that it automatically satisfies Expression (2.45). The exact theorems and proofs are given in Appendix B. In the literature we have also found proofs for similar methods. It is not clear to the author if the proofs found in [DF93], which are put in a very general setting, also apply to the forward and backward aggregation method that we described here. Though of mathematical importance, the proofs that we found can be skipped and are not necessary for the remainder of the thesis.

Remark. In this section we have followed a decomposition procedure based on two-machine production line models in which the machines have *geometrically* distributed service times. The decomposition procedure in this section, including the forward and backward aggregation procedures, is not new since it has been used for similar manufacturing systems with continuous service times in the past. It corresponds for instance to the procedure mentioned in [HB67] and many following articles in which we can find a similar method for servers with exponentional or other continuous service times. The method here is new in the sense that we cannot find preceding articles with applications to serial production lines with machines that have *discrete* geometrically distributed service times.

2.5 Synthesis: the Minimum Workforce Problem

Consider a serial production line with product failures only, of which we described the properties already in Subsection 2.1. Machine M_i has completion probability p_i, and buffer B_i with capacity N_i is located between machines M_{i-1} and M_i. In the previous section we described an algorithm to approximate the average production rate R of the total production line. The completion probability of a machine is related to the effort of the operator(s) working on it. So the completion probability of a particular machine can be seen as a measure for the workforce. We can therefore define the total workforce W as the sum of the completion probabilities:

$$W \stackrel{\text{def}}{=} \sum_{i=1}^{m} p_i. \tag{2.52}$$

The question in this subsection is: is it possible to obtain the same total production rate R, with the same buffer capacities N_i, but with less workforce? To examine this we will first approach the problem on the two-machine production line.

2.5.1 The Minimum Workforce Problem for a Two-Machine Production Line

We focus on two-machine production lines to examine again the average production rate R_N of the system as a function of the completion probabilities p_1 and p_2 of the machines. The relationship between p_1, p_2 and the average production rate R_N in case of buffer capacities $N = 1, 2, 3, 4$ are again shown in Figures 2.21, 2.22, 2.23 and 2.24. The figures correspond to the figures that we have already shown in Figures 2.10 up to 2.13. The difference with Figures 2.10 to 2.13 is that in Figures 2.21 to 2.24 we have drawn lines with equal workforce. From these figures we can see clearly that for every buffer capacity N on each "iso-workforce" line we obtain the highest average production rate if we take equal completion probabilities for both machines: $p_1 = p_2$. This can also be proved analytically.

Suppose we have a constant total workforce W and a first machine with completion probability x. This implies that the second machine has a completion probability $W - x$. It is our aim to find for which x the average production rate $R_N(x, W - x)$ takes its maximum value. To find the maximum we differentiate $R_N(x, W - x)$ to the variable x and get:

$$\frac{d}{dx}\{R_N(x, W - x)\} = \frac{\partial R_N}{\partial x}(x, W - x) - \frac{\partial R_N}{\partial y}(x, W - x).$$

Now we can use "R_N property 11" in Theorem 16 from Appendix A in Section A.2 several times.

If $x > \frac{1}{2}\,W$ then $x > W - x$ and from "R_N property 11" follows

$$\begin{aligned} \frac{\partial R_N}{\partial x}(x, W-x) &> \frac{\partial R_N}{\partial y}(x, W-x), \\ \frac{\partial R_N}{\partial x}(x, W-x) - \frac{\partial R_N}{\partial y}(x, W-x) &> 0, \\ \frac{d}{dx}\{R_N(x, W-x)\} &> 0. \end{aligned}$$

If $x < \frac{1}{2}\,W$ then $x < W - x$ and from "R_N property 11" follows

$$\begin{aligned} \frac{\partial R_N}{\partial x}(x, W-x) &< \frac{\partial R_N}{\partial y}(x, W-x), \\ \frac{\partial R_N}{\partial x}(x, W-x) - \frac{\partial R_N}{\partial y}(x, W-x) &< 0, \\ \frac{d}{dx}\{R_N(x, W-x)\} &< 0. \end{aligned}$$

This can only imply that the maximum for $R_N(x, W-x)$ can be obtained for $x = \frac{1}{2}\,W$ which means that for a maximal average production rate with a constant workforce both machines must have equal completion probabilities. Now we have found that the maximal average production rate R for a fixed workforce W is obtained when we take equal completion probabilities p_1 and p_2. As a consequence we can conclude that the minimal workforce W for a fixed average production rate R is obtained for equal completion probabilities p_1 and p_2 as well.

2.5.2 The Minimum Workforce Problem for a Multi-Machine Production Line

Is it possible to apply this result of the two-machine production line to systems with more machines? The answer to this question is quite simple if we realize that we already made a decomposition of a multi-machine production line into several two-machine production line models. For the minimum workforce problem of the multi-machine production line we just require that the all two-machine production models from the decomposition have equal completion probabilities. In other words, for all $i \in \{2,3,...,m\}$ we require that $p_{i-1}^f = p_i^b$. Hence for the minimal workforce the following relation holds:

$$\hat{p}_i \stackrel{\text{def}}{=} p_{i-1}^f = p_i^b, \quad \forall\, i \in \{2,3,..,m\}. \tag{2.53}$$

We can also derive the production rate of the two-machine production line model for buffer B_i.

$$R = R_{N_i}(\hat{p}_i, \hat{p}_i). \tag{2.54}$$

From "R_N property 2" in Theorem 7 from Appendix A in Section A.2 we derive that

$$R_{N_i}(\hat{p}_i, \hat{p}_i) = \frac{N_i \hat{p}_i}{N_i + 1 - \hat{p}_i}. \tag{2.55}$$

From (2.54) and (2.55) we derive:

$$\hat{p}_i = \frac{N_i + 1}{N_i + R} \cdot R. \tag{2.56}$$

From (2.48) and (2.53) follows:

$$\hat{p}_i = p_i.(1 - \hat{q}_{i+1}\Psi_{N_{i+1}}(\hat{q}_{i+1}, \hat{q}_{i+1})). \tag{2.57}$$

When we combine this with

$$\Psi_{N_{i+1}}(\hat{q}_{i+1}, \hat{q}_{i+1}) = \frac{1}{N_{i+1} + \hat{q}_{i+1}},$$

then we conclude from (2.57) that

$$\hat{p}_i = \frac{N_{i+1} p_i}{N_{i+1} + 1 - \hat{p}_{i+1}}. \tag{2.58}$$

From (2.56) and (2.58) we can finally derive a remarkable *balance equation*:

$$p_i = \frac{N_i + 1}{N_i + R} \cdot \frac{N_{i+1} + 1}{N_{i+1} + R} \cdot R, \tag{2.59}$$

$$\forall\, i \in \{1,2,..,m\} \quad N_1 \stackrel{\text{def}}{=} \infty, \quad N_{m+1} \stackrel{\text{def}}{=} \infty.$$

The answer to the minimum workforce problem is now solved. To obtain a required production rate R with minimum workforce we have to manage the completion probabilities of the machines so that (2.59) is satisfied. A production line that satisfies (2.59) will be called *dynamically balanced.*

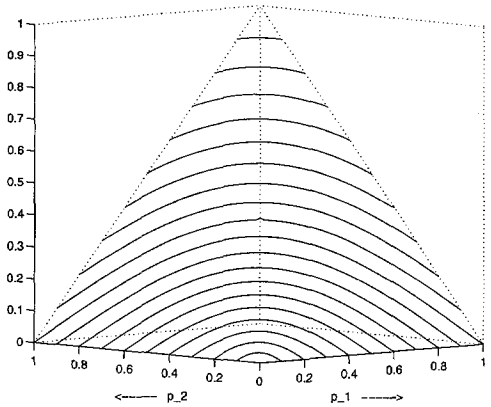

Figure 2.21: Three-Dimensional plot of R_1.

Figure 2.22: Three-Dimensional plot of R_2.

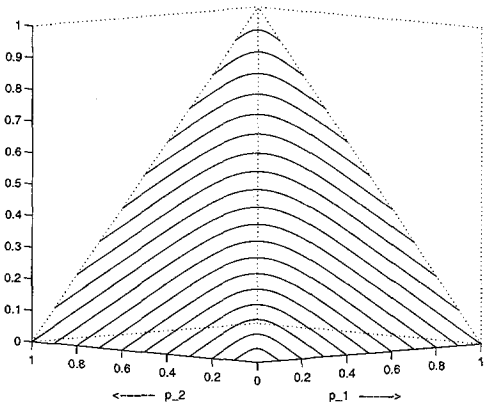

Figure 2.23: Three-Dimensional plot of R_3.

Figure 2.24: Three-Dimensional plot of R_4.

2.6 Synthesis: the Buffer Allocation Problem

In the previous subsection we presumed that the capacity of each buffer B_i was fixed to be N_i. The subsection resulted in an optimal distribution of the total workforce in separate completion probabilities as in (2.59). In this subsection we therefore examine the class of dynamically balanced production lines. We want to continue the optimization to find the best distribution of buffer capacities over the different buffers B_i. It is possible to approach this problem in two different ways:

1. We have $m-1$ buffers. Each buffer has a fixed capacity. What is the optimal order of succession of these buffers in the production line?
2. We have a total buffer capacity C. What is the optimal distribution of the total capacity C over the different buffers?

In both approaches the buffer capacities are assumed to be natural numbers.

Since we optimize within the class of dynamically balanced systems, (2.59) holds. The total workforce, as a function of the production rate R and buffer capacities N_i is therefore:

$$W(R, N_2, N_3, .., N_m) = R.\left(\sum_{i=1}^{m} \frac{N_i+1}{N_i+R} \cdot \frac{N_{i+1}+1}{N_{i+1}+R}\right). \tag{2.60}$$

Now we can reformulate the questions 2.6 and 2.6 as follows:

Problem 1: Given $m-1$ fixed buffer capacities $C_2 \leq C_3 \leq C_4 \leq ... \leq C_m$. What permutation $N_i = C_j$ of these capacities minimizes the function W in (2.60)?

Problem 2: Given a total buffer capacity C. Find the buffer capacities N_i that minimize function (2.60) with the constraint that $\sum_i N_i = C$.

We approached both problems with the help of a computer. We made a program that generated all possible permutations and found the solutions with specific parameters. All the results from the computations were in accordance with the following:

Solution to problem 1: Independent of the average production rate R, the optimal permutation for the minimization of the workforce in case $m > 9$ is:

$$\begin{array}{lll}
C_2C_m,\ C_4C_{m-2},.. & ,\ C_{\frac{m+2}{2}},.. & ,\ C_{m-3}C_5,\ C_{m-1}C_3,\ \text{if } m \bmod 4 = 2,\\
C_2C_m,\ C_4C_{m-2},.. & ,\ C_{\frac{m+1}{2}}C_{\frac{m+3}{2}},.. & ,\ C_{m-3}C_5,\ C_{m-1}C_3,\ \text{if } m \bmod 4 = 3,\\
C_2C_m,\ C_4C_{m-2},.. & ,\ C_{\frac{m}{2}}C_{\frac{m+4}{2}},\ C_{\frac{m+2}{2}},.. & ,\ C_{m-3}C_5,\ C_{m-1}C_3,\ \text{if } m \bmod 4 = 0,\\
C_2C_m,\ C_4C_{m-2},.. & ,\ C_{\frac{m-1}{2}}C_{\frac{m+5}{2}},\ C_{\frac{m+3}{2}}C_{\frac{m+1}{2}},.. & ,\ C_{m-3}C_5,\ C_{m-1}C_3,\ \text{if } m \bmod 4 = 1.
\end{array} \tag{2.61}$$

An intuitive explanation for this procedure is that the larger a buffer the better it can compensate the drawback of a small buffer next to it. Therefore it is reasonable that it is best to put the smallest capacity (which is the biggest drawback) next to the largest capacity (which is the best compensator). In fact we start with two buffers with infinite contents: $N_1 = \infty$ and $N_{m+1} = \infty$. Next to these infinite capacities, which are the best compensators, we put the smallest capacities, which are the worst bottlenecks. $N_2 = C_2$ and $N_m = C_3$. Next to the smallest capacity C_2 we put the largest capacity $N_3 = C_m$. Next to the second smallest capacity C_3 we put the second largest capacity $N_{m-1} = C_{m-1}$. If we continue this procedure of putting capacities from both ends to the middle, we will obtain the same permutation as the best permutation stated in Expression (2.61). From (2.60) we can see that reversing the order of the capacities will not change the workforce. Therefore the reverse order of the permutation of capacities that we have found here will also be an optimal solution.

Solution to problem 2: Cut the total buffer capacity C in $m-1$ separate buffer capacities. It is best to divide the total capacity in equal parts. Since the buffer capacities are natural numbers, it is not always possible to equalize all separate capacities, but it is possible to permit a maximal difference of 1 product part between any two capacities. If not all capacities are equal, then apply the algorithm described in Solution to problem 1.

2.7 The Average Buffer Contents in a Dynamically Balanced System

Theorem 1 *The average steady state buffer contents in a dynamically balanced production line is:*

$$\mathbb{E}\ (\text{Contents } B_i) = \frac{N_i + R}{2}, \quad \forall\, i \in \{1, 2, .., m-1\}, \tag{2.62}$$

where R is the average production rate of the production line.

Proof. With (2.14) and (2.56) follows:

$$\begin{aligned}
\mathbb{E}\ (\text{Contents } B_i) &= \sum_{j=0}^{N_i} j.\mathbb{P}\text{r}(\text{Contents } B_i = j) \\
&= \sum_{j=1}^{N_i} \frac{j}{N_i + \hat{q}_i} \\
&= \frac{1}{N_i + \hat{q}_i} \cdot \sum_{j=1}^{N_i} j
\end{aligned}$$

Buffer	Capacity	Buffer	Capacity
B_1	∞	B_{12}	11
B_2	1	B_{13}	10
B_3	20	B_{14}	13
B_4	3	B_{15}	8
B_5	18	B_{16}	15
B_6	5	B_{17}	6
B_7	16	B_{18}	17
B_8	7	B_{19}	4
B_9	14	B_{20}	19
B_{10}	9	B_{21}	2
B_{11}	12	B_{22}	∞

Table 2.1: The best buffer allocation to obtain minimal workforce.

$$= \frac{1}{N_i + (1 - \frac{N_i + 1}{N_i + R} R)} \cdot \frac{N_i(N_i + 1)}{2}$$

$$= \frac{N_i + R}{2}.$$

■

2.8 Example and Evaluation

Suppose that we have a serial production line of 21 machines and 20 buffers. The production line satisfies all assumptions mentioned in Section 2.1. Suppose also that the buffer capacities are fixed to be the natural numbers from 1 to 20: $C_i = i - 1, \quad \forall i \in \{2, 3, .., 21\}$. The allocation of the buffer capacities is to be chosen. Let us assume that the required average production rate for the total production line is restricted to be at least 0.5 product part per time slot. When we synthesize the production line the following questions arise:

a. What is the best arrangement of the buffer capacities in the line?

b. What is the minimal workforce to meet the requirements?

c. What completion probabilities correspond to the minimal workforce?

For the answer to question **a** we have to apply the result mentioned in Expression (2.61) from the previous subsection. This expression results in the best permutation of the buffer capacities shown in Table 2.1. For the answer to question **c** we have to make the system dynamically balanced. Therefore we can find the best completion probabilities by means of

Compl. prob.	value	approx.	Compl. prob.	value	approx.
p_1	$\frac{2}{3}$	.6667			
p_2	$\frac{84}{123}$	.6829	p_{12}	$\frac{264}{483}$	.5466
p_3	$\frac{168}{287}$	.5854	p_{13}	$\frac{308}{567}$	.5432
p_4	$\frac{152}{259}$	.5869	p_{14}	$\frac{252}{459}$	.5490
p_5	$\frac{228}{407}$	.5602	p_{15}	$\frac{288}{527}$	.5465
p_6	$\frac{204}{363}$	.5620	p_{16}	$\frac{224}{403}$	.5558
p_7	$\frac{272}{495}$	.5495	p_{17}	$\frac{252}{455}$	.5538
p_8	$\frac{240}{435}$	.5517	p_{18}	$\frac{180}{315}$	.5714
p_9	$\frac{300}{551}$	.5445	p_{19}	$\frac{200}{351}$	.5698
p_{10}	$\frac{260}{475}$	.5474	p_{20}	$\frac{120}{195}$	.6154
p_{11}	$\frac{312}{575}$	.5426	p_{21}	$\frac{3}{5}$	.6000

Table 2.2: The completion probabilities to obtain a dynamically balanced system.

Expression (2.59). Using (2.59) we obtain the completion probabilities in Table 2.2. In this table we can observe a phenomenon that is known as the *bowl phenomenon* in the literature (See [Rao76]). This phenomenon says that the lowest completion probabilities are in the central stages and increase progressively towards either side of the line. From Table 2.2 we can derive the answer to question **b**: the minimal total workforce $W = \sum_i p_i \approx 12.0313$. The minimal average workforce $W/m \approx \frac{12.0313}{21} \approx 0.5729$.

Since the formulas in this section are based on approximations, we want to evaluate the accuracy of the previous solution. Therefore we made a program for simulation of our production line in C programming language. The most important tool for the simulations is the random generator used. For our simulations in C we used the *drand48* command which according to the manual pages generates pseudo-random numbers using a well-known linear congruential algorithm and 48-bit integer arithmetic. The pseudo-random number generator is initialized by means of a time dependent argument so as to generate an independent sequence each time it is invoked.

The structure of the program is as follows. First we simulate the production line for one million time slots in order to estimate the average buffer contents of the separate buffers. Next we execute 10 independent runs of the production line, one million time slots per run. The buffer contents at the beginning of each of these runs is taken such that it is close to the overall average buffer contents. This is to get rid of the influence that the transient behaviour has on the results for the stationary behaviour of the system. Each of the 10 independent runs ends with an estimate of the average production rate R, and an estimate of the average buffer contents of the separate buffers. In this manner we have found 10

Estimates of average production rate R by simulation										
	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
R	0.494148	0.494314	0.494651	0.494199	0.494239	0.494575	0.494705	0.494403	0.494400	0.494200
Estimates of the average buffer contents by simulation										
Buffer	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
B_2	0.7529	0.7524	0.7536	0.7527	0.7521	0.7527	0.7524	0.7528	0.7527	0.7527
B_3	11.9460	11.7575	12.1317	11.9803	11.9498	12.0927	11.7341	11.7967	11.8360	11.8487
B_4	1.8004	1.8000	1.8012	1.8024	1.8072	1.7999	1.7889	1.7996	1.7917	1.7950
B_5	10.3001	10.4912	10.6484	10.6032	10.6194	10.7129	10.2754	10.4872	10.2054	10.3577
B_6	2.8446	2.8550	2.8772	2.8607	2.8746	2.8740	2.8356	2.8407	2.8500	2.8467
B_7	8.8378	8.9662	9.0997	9.0966	9.1252	9.0747	8.8213	8.8959	9.0804	8.8871
B_8	3.8508	3.8843	3.8971	3.8920	3.9159	3.8927	3.8557	3.9057	3.8920	3.8482
B_9	7.6536	7.6016	7.6100	7.5340	7.7646	7.5009	7.5015	7.6491	7.4747	7.4825
B_{10}	4.8670	4.8932	4.9247	4.8404	4.9445	4.8639	4.8515	4.8281	4.8220	4.8061
B_{11}	6.2793	6.3186	6.4725	6.3563	6.4396	6.3045	6.3273	6.3042	6.3429	6.2486
B_{12}	5.7146	5.7751	5.8804	5.7549	5.7589	5.7879	5.7741	5.7475	5.7044	5.6749
B_{13}	5.0797	5.1860	5.2800	5.2267	5.1714	5.1878	5.1442	5.1867	5.1573	5.0745
B_{14}	6.4708	6.6435	6.7272	6.6682	6.6134	6.6042	6.5135	6.5750	6.5987	6.4073
B_{15}	4.0695	4.1562	4.1756	4.1636	4.1297	4.1132	4.1147	4.1022	4.1786	4.0965
B_{16}	7.1769	7.3221	7.3946	7.3381	7.2222	7.2582	7.2976	7.1730	7.5197	7.1957
B_{17}	3.1264	3.1122	3.1635	3.1368	3.1005	3.1179	3.1330	3.1136	3.1747	3.1364
B_{18}	7.8876	7.7368	8.0619	7.7668	7.8088	7.9237	7.8038	7.8810	8.0231	7.7920
B_{19}	2.1628	2.1658	2.1843	2.1714	2.1764	2.1782	2.1690	2.1684	2.1809	2.1684
B_{20}	8.4867	8.3243	8.8084	8.3466	8.5155	8.4995	8.3785	8.4214	8.8619	8.4074
B_{21}	1.2320	1.2308	1.2349	1.2308	1.2330	1.2315	1.2310	1.2343	1.2337	1.2318

Table 2.3: The results of 10^6 time steps simulations for the 21-machine-production line

independent estimates of the average production rate and the corresponding average buffer contents for each buffer. This procedure allows us to do some simple statistical analysis such as estimate of mean value and variance. The direct simulation results are shown in Table 2.3. The computed mean values and standard deviations are shown in Table 2.4. The results in Table 2.4 are quite representative for the performance of the decomposition method presented here. The estimate of the average production rate R by means of the decomposition method usually stays within a range of 2 % relative error (1.14 % in our example), while the estimate of the average buffer contents can reach a relative error of more than 10 % (up to 14.6 % for B_{20} in our example).

□

2.9 Conclusions

In this chapter we described a decomposition method for a discrete synchronous serial production line with product failures only. Such a production line is in fact a serial queueing network with blocking in which service times are geometrically distributed. We used the decomposition method to derive the stationary distribution and the average production

Results for the average production rate R					
	mean value	std. deviation	decomp. model	abs. err.	rel. err.
R	0.494383	0.000200	0.500000	0.005617	0.0114

Results for the average buffer contents					
Buffer	mean value	std. deviation	decomp. model	abs. err.	rel. err.
B_2	0.7527	0.0004	0.7500	-0.0027	-0.0036
B_3	11.9073	0.1360	10.2500	-1.6573	-0.1391
B_4	1.7986	0.0053	1.7500	-0.0486	-0.0270
B_5	10.4701	0.1767	9.2500	-1.2201	-0.1165
B_6	2.8559	0.0151	2.7500	-0.1059	-0.0371
B_7	8.9885	0.1195	8.2500	-0.7385	-0.0821
B_8	3.8834	0.0237	3.7500	-0.1334	-0.0344
B_9	7.5772	0.0947	7.2500	-0.3272	-0.0432
B_{10}	4.8641	0.0449	4.7500	-0.1141	-0.0235
B_{11}	6.3394	0.0691	6.2500	-0.0894	-0.0141
B_{12}	5.7573	0.0560	5.7500	-0.0073	-0.0013
B_{13}	5.1694	0.0617	5.2500	0.0806	0.0156
B_{14}	6.5822	0.0953	6.7500	0.1678	0.0255
B_{15}	4.1300	0.0370	4.2500	0.1200	0.0291
B_{16}	7.2898	0.1096	7.7500	0.4602	0.0631
B_{17}	3.1315	0.0231	3.2500	0.1185	0.0378
B_{18}	7.8685	0.1086	8.7500	0.8815	0.1120
B_{19}	2.1726	0.0070	2.2500	0.0774	0.0356
B_{20}	8.5050	0.1856	9.7500	1.2450	0.1464
B_{21}	1.2324	0.0015	1.2500	0.0176	0.0143

Table 2.4: The compound results of the simulations for the 21-machine-production line

rate R of such production lines. The method consists of an algorithm with alternate forward and backward aggregation procedures. Similar algorithms have been used for queueing networks with blocking for about thirty years and can be considered classic. Nevertheless we think that the method described in this chapter is new, since preceding applications of these classic concepts to queueing networks with geometric service times have not been found. We proved convergence, existence, uniqueness and conservation of flow of the solution of the algorithm for every possible combination of parameters in the line.

Moreover, as a result from the decomposition method, we could easily derive properties for an optimally balanced production line, including a method for the optimization of the placement of buffers. These properties of optimallly balanced production lines appear to be in accordance with the "bowl phenomenon" described in the literature.

In order to evaluate the method we compared its results with results from simulations. It turned out that the estimate of the average production rate R usually was within a range of 2 % relative error, which can be considered as quite accurate. The estimate of the average buffer contents however could be worse, up to more than 10 % relative error. The big advantage of the method is that it takes very few computations to find the stationary distribution for the decomposed production line. An average PC can do all computations within seconds.

Chapter 3

Serial Production Lines with Machine Failures Only

In the previous chapter we have examined serial production lines with product failures only, which in fact are serial production lines in which machines have geometrically distributed service times. In this chapter we will drop the geometically distributed service times and try to describe machines that have constant service times. Instead we introduce new stochastic phenomena: machines that are prone to breakdown and repair. Therefore in this chapter we examine serial production lines with machine failures only. Before we examine this kind of production line we will first describe how such production lines work.

3.1 The Description of a Production Line with Machine Failures Only

The dynamics of the production line with machine failures only that we will describe here are basically the same as the dynamics of a production line with product failures only described in Section 2.1. The difference between the model described in Section 2.1 and the model we use for production lines with machine failures only is as follows. Since we assume here that product failures do not occur, we assume that the completion probabilities p_i of all machines M_i, $i \in \{1, 2, ..., m\}$ switch between 0 and 1. A completion probability $p_i = 0$ represents the stand still of a machine because of a machine failure. A completion probability $p_i = 1$ represents the completion probability of a machine in progress without product failures. In contrast with the production lines with product failures only, here we do *not* assume that completion probabilities p_i are constant in time. Therefore we assume that the completion probabilities p_i alternately change in time from 1 to 0 and back from 0 to 1. The "state" of a machine in which the completion probability equals 1 is called *up*. The state in which the completion probability equals 0 is called *down*. The transition from state up to state down is called *breakdown*. The transition from state down to state up is

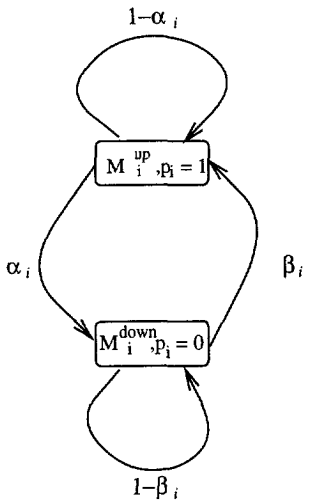

Figure 3.1: The breakdown-repair model of a machine M_i.

called *repair*. The transitions for all machines M_i between up and down are governed by a *machine breakdown-repair model.*

3.2 The Machine Breakdown-Repair Model

The machine breakdown-repair model is very simple: We assume that a machine M_i can only be in two different states: it is either up or down.

When up, the machine has a certain probability to go down during the current time slot. This probability is called the *breakdown rate* α_i. If the machine does not break down in the current time slot it will be up when the next time slot starts. When the machine does break down during the current time slot, it will be down at the beginning of the next time slot.
When down, the machine has a certain probability to go up during the current time slot. This probability is called the *repair rate* β_i. If the machine will not be repaired during the current time slot it will be down when the next time slot starts. When the machine is repaired during the current time slot, it will be up at the beginning of the next time slot.

An illustration of the breakdown-repair model is shown in Figure 3.1.

The breakdown-repair process for one single machine M_i can be described by the following

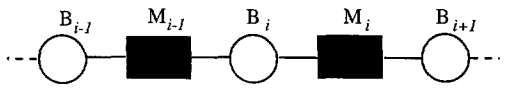

Figure 3.2: A part of the production line, focused on buffer B_i.

simple Markov chain:

$$\pi(t+1) = \begin{pmatrix} 1-\alpha_i & \beta_i \\ \alpha_i & 1-\beta_i \end{pmatrix} \pi(t) \ . \tag{3.1}$$

The stationary distribution π of this Markov chain is:

$$\pi = \frac{1}{\alpha_i + \beta_i} \begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix} . \tag{3.2}$$

Now the breakdown-repair model of the separate machines has been fully explained.

3.3 The Model for Buffers: the Equivalent of the Two-Machine Production Line Model

We now examine the buffer behaviour closely. For examining this behaviour we focus on a buffer B_i from the production line together with two surrounding machines and buffers. In this manner we get an arbitrary piece of a production line as shown in Figure 3.2. We will derive a Markov chain for buffer B_i by means of Figure 3.2. We therefore distinguish between buffer B_i and its environment. The *environment of buffer* B_i consists of two neighbouring machines M_{i-1} and M_i and two neighbouring buffers B_{i-1} and B_{i+1}.

3.4 The State Set of a Buffer

We focus on buffer B_i. The contents of this buffer be any natural number from 0 up to N_i. We closely examine what determines the changes in the buffer contents. What causes the buffer contents to increase, remain unchanged or decrease after a transition? Is it the current contents of the buffer that determines an increment or decrement after a transition, or is there more? First of all, the contents of buffer B_i after a transition is determined by:

- The contents of buffer B_i before the transition.

It turns out that it is not the contents of buffer B_i only that determines the contents of this buffer after a transition. Other important factors that determine the contents of buffer B_i after a transition are:

- Is the previous buffer B_{i-1} starved?

- Is the previous machine M_{i-1} down?
- Is the next machine M_i down?
- Is the next buffer B_{i+1} blocked?

The five items above determine the situation, the state that buffer B_i is in at a specific moment. These five items determine the state set of the buffer, they determine the state set of the Markov chain corresponding to buffer B_i. We will denote the Markov chain corresponding to buffer B_i by the term *two-machine Markov model for* B_i. If we combine the previous observations and define the sets

$$\begin{aligned} \mathcal{S}^{\text{Contents } B_i} &\stackrel{\text{def}}{=} \{0, 1, .., N_i\} , \\ \mathcal{S}^{\text{Blocking } B_{i+1}} &\stackrel{\text{def}}{=} \{B_{i+1}^{\neg\text{bl}}, B_{i+1}^{\text{bl}}\} , \\ \mathcal{S}^{\text{Starving } B_{i-1}} &\stackrel{\text{def}}{=} \{B_{i-1}^{\neg\text{st}}, B_{i-1}^{\text{st}}\} , \\ \mathcal{S}^{\text{State } M_j} &\stackrel{\text{def}}{=} \{M_j^{\text{up}}, M_j^{\text{down}}\} , \quad j = 1,\ j = i-1, \end{aligned} \tag{3.3}$$

we can define the total state set of buffer B_i, the state set of the two-machine Markov model for B_i, as follows:

$$\mathcal{S}^{B_i} \stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{i-1}} \times \mathcal{S}^{\text{State } M_{i-1}} \times \mathcal{S}^{\text{Contents } B_i} \times \mathcal{S}^{\text{State } M_i} \times \mathcal{S}^{\text{Blocking } B_{i+1}} , \tag{3.4}$$

where $\times$ stands for the Cartesian product of sets.
In the above definitions $B_i^{\neg\text{bl}}$, B_i^{bl}, $B_i^{\neg\text{st}}$ and B_i^{st} are used. These are defined as follows:

$$\begin{aligned} B_i^{\text{bl}} &\stackrel{\text{def}}{=} \{(a,b,c,d,e) \in \mathcal{S}^{B_i} \mid c = N_i \wedge (d = M_i^{\text{down}} \vee e = B_{i+1}^{\text{bl}})\} \\ B_i^{\neg\text{bl}} &\stackrel{\text{def}}{=} \{(a,b,c,d,e) \in \mathcal{S}^{B_i} \mid (a,b,c,d,e) \notin B_i^{\text{bl}}\} \\ B_i^{\text{st}} &\stackrel{\text{def}}{=} \{(a,b,c,d,e) \in \mathcal{S}^{B_i} \mid c = 0 \wedge (b = M_{i-1}^{\text{down}} \vee a = B_{i-1}^{\text{st}})\} \\ B_i^{\neg\text{st}} &\stackrel{\text{def}}{=} \{(a,b,c,d,e) \in \mathcal{S}^{B_i} \mid (a,b,c,d,e) \notin B_i^{\text{st}}\} \end{aligned} \tag{3.5}$$

Notice that the above definitions are of a recursive nature. The structure of the recursion is important for understanding the links between the buffer models. It justifies the recursive numerical aggregation method chosen in one of the next sections. Notice also that

$$\begin{aligned} B_i^{\text{bl}} &\cap B_i^{\text{st}} = \phi , \\ B_i^{\neg\text{bl}} &\cap B_i^{\neg\text{st}} \neq \phi , \\ B_i^{\text{bl}} &\subset B_i^{\neg\text{st}} , \\ B_i^{\text{st}} &\subset B_i^{\neg\text{bl}} . \end{aligned}$$

With the definition of the state set $\mathcal{S}^{B_i}$ of buffer B_i in mind we can, for a fixed contents c of the buffer, distinguish 16 different states in the state set. A table of these states is

State in $\mathcal{R}^{B_i}$	Corresponding state in $\mathcal{S}^{B_i}$				
	B_{i-1}	M_{i-1}	B_i	M_i	B_{i+1}
(c,1)	¬starved	up	c	up	¬blocked
(c,2)	¬starved	up	c	up	blocked
(c,3)	¬starved	up	c	down	¬blocked
(c,4)	¬starved	up	c	down	blocked
(c,5)	starved	up	c	up	¬blocked
(c,6)	starved	up	c	up	blocked
(c,7)	starved	up	c	down	¬blocked
(c,8)	starved	up	c	down	blocked
(c,9)	¬starved	down	c	up	¬blocked
(c,10)	¬starved	down	c	up	blocked
(c,11)	¬starved	down	c	down	¬blocked
(c,12)	¬starved	down	c	down	blocked
(c,13)	starved	down	c	up	¬blocked
(c,14)	starved	down	c	up	blocked
(c,15)	starved	down	c	down	¬blocked
(c,16)	starved	down	c	down	blocked

Table 3.1: The one to one correspondence of $\mathcal{S}^{B_i}$ and $\mathcal{R}^{B_i}$.

shown in Table 3.1. In this table we redefine the state set in another form. The new form simplifies the notation in future expressions. In the new definition a state is a combination of two numbers, the first number determines the contents of the corresponding buffer, the second number determines the state of the surrounding buffers and machines also denoted by the term *environment*. Each of the 16 possible combinations of states of the environment has as specific number listed in Table 3.1. The new state set $\mathcal{R}^{B_i}$, that has a one to one correspondence to the old state set $\mathcal{S}^{B_i}$, is defined as:

$$\mathcal{R}^{B_i} \stackrel{\text{def}}{=} \{0,\ 1,\ 2,..\ ,\ N_i\} \times \{1,\ 2,..\ \ ,\ 16\} \tag{3.6}$$

In special circumstances when $c = 0$ or $c = N_i$ in Table 3.1 a conflict can occur:

- When $c = 0$ then it can happen that M_{i-1} puts a part in B_i and at the same time M_i attempts to take a part from B_i. This only occurs in conflict situation $(0, 1) \in \mathcal{R}^{B_i}$.
- When $c = N_i$ then it can happen that M_{i-1} attempts to put a part in B_i and at the same time M_i takes a part from B_i. This only occurs in conflict situation $(N_i, 1) \in \mathcal{R}^{B_i}$.

In these situations the order in which machines M_{i-1} and M_i put and take a part is crucial for the resulting transition. In case of such conflicts we impose that within one time cycle machine M_i attempts to take a part *before* M_{i-1} can put one.
Having imposed this order we can conclude in accordance with definitions (3.5) that buffer B_i is starved in all cases with $c = 0$ in Table 3.1. Buffer B_i is blocked in case of a combination of $c = N_i$ with states $2 - 4$, $6 - 8$, $10 - 12$ and $14 - 16$ in Table 3.1. The

combination of $c = N_i$ and 1,5,9 or 13 are *not* considered to be blocked states. This is because in these states machine M_i will first take a product from buffer B_i and reduce the number of parts in it to $N_i - 1$. Next machine M_{i-1} "has the opportunity" to put a part in the last position of the buffer and therefore it can not be considered blocked. In total we have $16(N_i + 1)$ states in the state set $\mathcal{R}^{B_i}$ for a single buffer B_i.

3.5 Useful Partitionings in the Set $\mathcal{P}(\mathcal{R}^{B_i})$

In total we consider four main and three intermediate partitionings of $\mathcal{R}^{B_i}$. The main partitionings all correspond to a combination of a subsequent machine and buffer. In order to be precise the partitionings are described in detail in this section. From the name of each subset it will be clear in the following to what situation it corresponds. See also Table 3.1.

The first main partitioning $\rho^{B_{i-1},M_{i-1}} \in \mathcal{P}(\mathcal{R}^{B_i})$ is as follows:

$$\rho_i^{B_{i-1},M_{i-1}} \stackrel{\text{def}}{=} (\mathcal{R}_i^{(B_{i-1}^{\neg\text{st}},M_{i-1}^{\text{up}})}, \mathcal{R}_i^{(B_{i-1}^{\text{st}},M_{i-1}^{\text{up}})}, \mathcal{R}_i^{(B_{i-1}^{\neg\text{st}},M_{i-1}^{\text{down}})}, \mathcal{R}_i^{(B_{i-1}^{\text{st}},M_{i-1}^{\text{down}})}),$$

where

$$\begin{aligned}
\mathcal{R}_i^{(B_{i-1}^{\neg\text{st}},M_{i-1}^{\text{up}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{1,2,3,4\} \,\}, \\
\mathcal{R}_i^{(B_{i-1}^{\text{st}},M_{i-1}^{\text{up}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{5,6,7,8\} \,\}, \\
\mathcal{R}_i^{(B_{i-1}^{\neg\text{st}},M_{i-1}^{\text{down}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{9,10,11,12\} \,\}, \\
\mathcal{R}_i^{(B_{i-1}^{\text{st}},M_{i-1}^{\text{down}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{13,14,15,16\} \,\}.
\end{aligned} \tag{3.7}$$

The second main partitioning $\rho_i^{M_i,B_{i+1}} \in \mathcal{P}(\mathcal{R}^{B_i})$ used is as follows:

$$\rho_i^{M_i,B_{i+1}} \stackrel{\text{def}}{=} (\mathcal{R}_i^{(M_i^{\text{up}},B_{i+1}^{\neg\text{bl}})}, \mathcal{R}_i^{(M_i^{\text{up}},B_{i+1}^{\text{bl}})}, \mathcal{R}_i^{(M_i^{\text{down}},B_{i+1}^{\neg\text{bl}})}, \mathcal{R}_i^{(M_i^{\text{down}},B_{i+1}^{\text{bl}})}),$$

where

$$\begin{aligned}
\mathcal{R}_i^{(M_i^{\text{up}},B_{i+1}^{\neg\text{bl}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{1,5,9,13\} \,\}, \\
\mathcal{R}_i^{(M_i^{\text{up}},B_{i+1}^{\text{bl}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{2,6,10,14\} \,\}, \\
\mathcal{R}_i^{(M_i^{\text{down}},B_{i+1}^{\neg\text{bl}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{3,7,11,15\} \,\}, \\
\mathcal{R}_i^{(M_i^{\text{down}},B_{i+1}^{\text{bl}})} &= \{\, (a,b) \in \mathcal{R}^{B_i} \mid b \in \{4,8,12,16\} \,\}.
\end{aligned} \tag{3.8}$$

The third main partitioning $\rho_i^{M_{i-1},B_i} \in \mathcal{P}(\mathcal{R}^{B_i})$ is:

$$\rho_i^{M_{i-1},B_i} \stackrel{\text{def}}{=} (\mathcal{R}_i^{(M_{i-1}^{\text{up}},B_i^{\neg\text{bl}})}, \mathcal{R}_i^{(M_{i-1}^{\text{up}},B_i^{\text{bl}})}, \mathcal{R}_i^{(M_{i-1}^{\text{down}},B_i^{\neg\text{bl}})}, \mathcal{R}_i^{(M_{i-1}^{\text{down}},B_i^{\text{bl}})}),$$

where

$$\begin{aligned}
\mathcal{R}_i^{(M_{i-1}^{\mathrm{up}},B_i^{\neg\mathrm{bl}})} &= \{(a,b)\in\mathcal{R}^{B_i} | (a<N_i \wedge b<9) \vee (a=N_i \wedge b\in\{1,5\})\},\\
\mathcal{R}_i^{(M_{i-1}^{\mathrm{up}},B_i^{\mathrm{bl}})} &= \{(a,b)\in\mathcal{R}^{B_i} | a=N_i \wedge b\in\{2,3,4,6,7,8\}\},\\
\mathcal{R}_i^{(M_{i-1}^{\mathrm{down}},B_i^{\neg\mathrm{bl}})} &= \{(a,b)\in\mathcal{R}^{B_i} | (a<N_i \wedge b>8) \vee (a=N_i \wedge b\in\{9,13\})\},\\
\mathcal{R}_i^{(M_{i-1}^{\mathrm{down}},B_i^{\mathrm{bl}})} &= \{(a,b)\in\mathcal{R}^{B_i} | a=N_i \wedge b\in\{10,11,12,14,15,16\}\}.
\end{aligned} \tag{3.9}$$

The fourth main partitioning $\rho_i^{B_i,M_i} \in \mathcal{P}(\mathcal{R}^{B_i})$ is:

$$\rho_i^{B_i,M_i} \stackrel{\mathrm{def}}{=} (\, \mathcal{R}_i^{(B_i^{\neg\mathrm{st}},M_i^{\mathrm{up}})}, \mathcal{R}_i^{(B_i^{\mathrm{st}},M_i^{\mathrm{up}})}, \mathcal{R}_i^{(B_i^{\neg\mathrm{st}},M_i^{\mathrm{down}})}, \mathcal{R}_i^{(B_i^{\mathrm{st}},M_i^{\mathrm{down}})} \,),$$

where

$$\begin{aligned}
\mathcal{R}_i^{(B_i^{\neg\mathrm{st}},M_i^{\mathrm{up}})} &= \{(a,b)\in\mathcal{R}^{B_i} | a>0 \wedge b\in\{1,2,5,6,9,10,13,14\}\},\\
\mathcal{R}_i^{(B_i^{\mathrm{st}},M_i^{\mathrm{up}})} &= \{(a,b)\in\mathcal{R}^{B_i} | a=0 \wedge b\in\{1,2,5,6,9,10,13,14\}\},\\
\mathcal{R}_i^{(B_i^{\neg\mathrm{st}},M_i^{\mathrm{down}})} &= \{(a,b)\in\mathcal{R}^{B_i} | a>0 \wedge b\in\{3,4,7,8,11,12,15,16\}\},\\
\mathcal{R}_i^{(B_i^{\mathrm{st}},M_i^{\mathrm{down}})} &= \{(a,b)\in\mathcal{R}^{B_i} | a=0 \wedge b\in\{3,4,7,8,11,12,15,16\}\}.
\end{aligned} \tag{3.10}$$

Next to the main partitionings we also use the following intermediate partitioning $\rho_i^{\mathrm{Env}} \in \mathcal{P}(\mathcal{R}^{B_i})$ of the environment of buffer B_i:

$$\rho_i^{\mathrm{Env}} \stackrel{\mathrm{def}}{=} (\, \mathcal{R}_i^{(\mathrm{Env},1)}, \mathcal{R}_i^{(\mathrm{Env},2)}, ..., \mathcal{R}_i^{(\mathrm{Env},16)} \,),$$

where

$$\mathcal{R}_i^{(\mathrm{Env},j)} = \{(a,b)\in\mathcal{R}^{B_i} | b=j\}. \tag{3.11}$$

Another intermediate partitioning $\rho_i^{\mathrm{Cont}} \in \mathcal{P}(\mathcal{R}^{B_i})$ corresponds directly to the contents of the buffer B_i:

$$\rho_i^{\mathrm{Cont}} \stackrel{\mathrm{def}}{=} (\, \mathcal{R}_i^{(\mathrm{Cont},0)}, \mathcal{R}_i^{(\mathrm{Cont},1)}, \mathcal{R}_i^{(\mathrm{Cont},2)}, ..., \mathcal{R}_i^{(\mathrm{Cont},N_i)} \,),$$

where

$$\mathcal{R}_i^{(\mathrm{Cont},j)} = \{(a,b)\in\mathcal{R}^{B_i} | a=j\}. \tag{3.12}$$

3.6 Construction of the Transition Matrix P^{B_i} of the Two-Machine Markov Model for Buffer B_i

In the previous subsections we have made a thorough description of the separate states and clusters of states that we can distinguish for the two-machine Markov model for buffer B_i. Determining the states of a Markov chain however is not enough for a Markov chain model description. The next step is to describe how the transition probabilities between

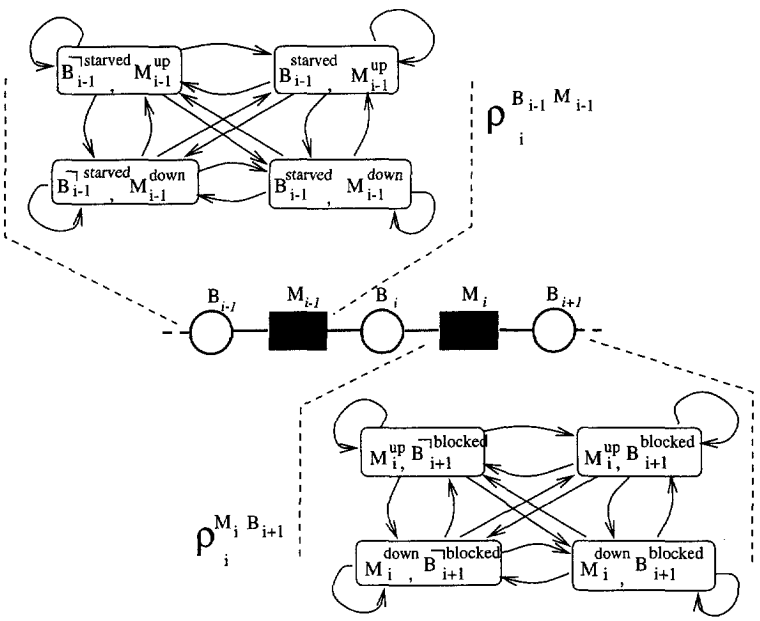

Figure 3.3: The two sub-chains for construction with partitionings $\rho_i^{B_{i-1},M_{i-1}}$ and $\rho_i^{M_i,B_{i+1}}$.

the different states should be determined. If we recall the construction of the Markov chain corresponding to a two-machine production line with product failures only described in Chapter 2, then we can say that there the Markov chain model is "driven" by the completion probabilities of the two machines in the model. These completion probabilities were considered to be independent. For the two-machine production lines with machine failures only that we describe here we will assume that the Markov chain model is "driven" by the Markov sub-chains corresponding to partitionings $\rho_i^{B_{i-1},M_{i-1}}$ and $\rho_i^{M_i,B_{i+1}}$ defined in Expressions (3.7) and (3.8). These two Markov chains again correspond to both machines M_{i-1} and M_i respectively. Again we will assume that the two processes that drive our model are independent processes. A picture of the two independent Markov chains that "drive" the two-machine Markov model for buffer B_i is shown in Figure 3.3.
For the construction of transition matrix P^{B_i} we have to determine how transitions between the states in $\mathcal{R}^{B_i}$ take place. We described two intermediate partitionings ρ_i^{Env} and ρ_i^{Cont} in the previous subsection. The transitions in partitioning ρ_i^{Env}, which consists of combinations from $\rho_i^{B_{i-1},M_{i-1}}$ and $\rho_i^{M_i,B_{i+1}}$, "drive" the two-machine Markov model and the transitions in partitioning ρ_i^{Cont} depend on the current state in ρ_i^{Env}.

In this chapter, as well as in the following two chapters, we will assume a specific order in the transitions. We will assume that one transition in $\mathcal{R}^{B_i}$ in one time slot is split in two subsequent transitions:

1. *Firstly* there will be a transition between clusters in partitioning ρ_i^{Cont} and

2. *Secondly* there will be a transition between clusters in partitioning ρ_i^{Env}.

The combination of both transitions determines the total transition in $\mathcal{R}^{B_i}$ for one time slot. In the next subsections we will:

- Firstly describe how we determine transitions in ρ_i^{Env}, described by means of the environment transition matrices $P_{\rho_i^{\text{Env}}}$ and $P^{B_i,\text{Env}}$ of which the last will be defined later.

- Next we will describe the relation between the state in ρ_i^{Env} and the resulting transition in ρ_i^{Cont}. This will result in a transition matrix for the transitions in contents denoted by $P^{B_i,\text{Cont}}$ defined later.

- After that we will be able to construct the whole transition matrix P^{B_i} of the two-machine Markov model for buffer B_i by means of simple multiplication of the two transition matrices $P^{B_i,\text{Env}}$ and $P^{B_i,\text{Cont}}$ that we obtained in the two previous steps.

3.6.1 The Construction of the Environment Transition Matrix $P_{\rho_i^{\text{Env}}}$ for Buffer B_i

In order to construct the transition matrix $P^{B_i,\text{env}}$ for the transitions in the environment of the two-machine Markov model for buffer B_i we first construct the Markov model for the environment of buffer B_i. For that purpose we will use the *Kronecker product.*
If two homogeneous discrete-time finite Markov chains Markov processes with transition matrices P_1 and P_2 are totally independent then the transition matrix P^{Comb} corresponding to all combinations of the states of both Markov processes can be obtained by means of the so called *Kronecker product* of matrices: $P^{\text{Comb}} = P_1 \otimes P_2$. This Kronecker product of two matrices A and B, denoted by "$A \otimes B$", is defined as follows:

$$A \otimes B \stackrel{\text{def}}{=} \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix} .$$

By the assumption that both Markov processes shown in Figure 3.3, which have transition matrices $P_{\rho_i^{B_{i-1},M_{i-1}}}$ and $P_{\rho_i^{M_i,B_{i+1}}}$ are independent we can now construct the transition matrix for the environment $P_{\rho_i^{\text{Env}}}$, which is the combined Markov process, by means of the Kronecker product:

$$P_{\rho_i^{\text{Env}}} = P_{\rho_i^{B_{i-1},M_{i-1}}} \otimes P_{\rho_i^{M_i,B_{i+1}}} . \tag{3.13}$$

By means of this transition matrix $P_{\rho_i^{\text{Env}}}$ we can easily construct the bigger transition matrix $P^{B_i,\text{env}}$ for the transitions in the environment of the two-machine Markov model for buffer B_i. This will be done in Subsection 3.6.4.

3.6.2 The Relation Between the Environment State in ρ_i^{Env} and the Transitions in ρ_i^{Cont}

In the previous subsection we described how to determine probabilities of the transitions in ρ_i^{Env}. In this subsection we assume that there are *no* transitions in ρ_i^{Env}, but only transitions in ρ_i^{Cont} induced by a fixed state in ρ_i^{Env}. A transition in ρ_i^{Cont} means that we start in a certain state $\mathcal{R}_i^{(\text{Cont,j})}$ and jump to another state $\mathcal{R}_i^{(\text{Cont,k})}$. (See also Expression (3.12).) In general there are only three possibilities:

- The contents of buffer B_i increases by one product after a transition.

$$k = j + 1,$$

- The contents of buffer B_i remains the same after a transition.

$$k = j,$$

- The contents of buffer B_i decreases by one product after a transition.

$$k = j - 1,$$

Of course not all three possibilities exist in case the contents of buffer B_i before the transition equals 0 or N_i since the contents of an empty buffer cannot decrease and the contents of a full buffer cannot increase. That is why these two cases are examined separately. In the next two paragraphs we will describe the three different cases for increment, no change, decrement separately:

The general case: starting position non-empty and non-full.

For the case that we start with a non-empty and non-full buffer, so in case of a starting position in $\bigcup_{j=1}^{N_i-1} \mathcal{R}_i^{(\text{Cont},j)}$, we have the following transitions:

- **Increment.**
  The contents of buffer B_i increases by one product if buffer B_{i-1} is not starved and machine M_{i-1} is up and either machine M_i is down or buffer B_{i+1} is blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{2,3,4\}\}.$$

  We translate this case in a corresponding diagonal matrix Λ_E as follows:

$$\Lambda_E \stackrel{\text{def}}{=} \text{diag}(0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0)$$

- **No change.**
  The contents of buffer B_i remains the same in two sub-cases

  - In the sub-case that buffer B_{i-1} is not starved and machine M_{i-1} is up and machine M_i is up and buffer B_{i+1} is not blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{1\}\}.$$

- In the sub-case that buffer B_{i-1} is starved or machine M_{i-1} is down and either machine M_i is down or buffer B_{i+1} is blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{6,7,8,10,11,12,14,15,16\}\}.$$

We translate both these sub-cases in a corresponding diagonal matrix Λ_D as follows:

$$\Lambda_D \stackrel{\text{def}}{=} \text{diag}(1,0,0,0,0,1,1,1,0,1,1,1,0,1,1,1)$$

- **Decrement.**
The contents of buffer B_i decreases by one product if buffer B_{i-1} is starved or machine M_{i-1} is down and both machine M_i is up and buffer B_{i+1} is not blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{5,9,13\}\}.$$

We translate this case in a corresponding diagonal matrix Λ_C as follows:

$$\Lambda_C \stackrel{\text{def}}{=} \text{diag}(0,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0)$$

The special cases: starting position empty or full.

Next we will describe the two separate cases in which we start with either an empty or a full buffer B_i, so the case of a starting position in $\mathcal{R}_i^{(\text{Cont},0)}$ or in $\mathcal{R}_i^{(\text{Cont},N_i)}$. In these cases we have to be very careful because they include the "conflict situations" $(0,1)$ and $(N_i,1)$ in $\mathcal{R}^{B_i}$ which we described earlier.

Empty starting position
First we will describe the case with an empty buffer B_i as a starting position, a starting position in $\mathcal{R}_i^{(\text{Cont},0)}$. In that special case we have the following transitions:

- **Increment.**
The contents of buffer B_i increases by one product in two sub-cases:

  - In the sub-case that buffer B_{i-1} is not starved and machine M_{i-1} is up and either machine M_i is down or buffer B_{i+1} is blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{2,3,4\}\}.$$

  - In the sub-case that we are in conflict situation $(0,1)$ in $\mathcal{R}^{B_i}$ which here corresponds to:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{1\}\}.$$

We translate both sub-cases in a corresponding diagonal matrix Λ_B as follows:

$$\Lambda_B \stackrel{\text{def}}{=} \text{diag}(1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0)$$

- **No change.**
  The contents of buffer B_i remains the same in all cases in which either buffer B_{i-1} is starved or machine M_{i-1} is down. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{5,6,7,8,9,10,11,12,13,14,15,16\}\}.$$

  We translate this case in a corresponding diagonal matrix Λ_A as follows:

$$\Lambda_A \stackrel{\text{def}}{=} \text{diag}(0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1)$$

- **Decrement.**
  Impossible.

Full starting position

Last we will describe the case with a full buffer B_i as a starting position, a starting position in $\mathcal{R}_i^{(\text{Cont},N_i)}$. In that special case we have the following transitions:

- **Increment.**
  Impossible.

- **No change.**
  The contents of buffer B_i remains the same in the following two sub-cases:

  - The sub-case in which either machine M_i is down or buffer B_{i+1} is blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{2,3,4,6,7,8,10,11,12,14,15,16\}\}.$$

  - The sub-case of conflict situation $(N_i, 1)$ in $\mathcal{R}^{B_i}$ which here corresponds to:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{1\}\}.$$

  We translate both sub-cases in a corresponding diagonal matrix Λ_G as follows:

$$\Lambda_G \stackrel{\text{def}}{=} \text{diag}(1,1,1,1,0,1,1,1,0,1,1,1,0,1,1,1)$$

  If we examine Λ_G and the diagonal matrices Λ_D and Λ_E that we defined earlier, then we can conclude that:

$$\Lambda_G = \Lambda_D + \Lambda_E,$$

- **Decrement.**
  The contents of buffer B_i decreases by one product in case buffer B_{i-1} is starved or machine M_{i-1} is down and both machine M_i is up and buffer B_{i+1} is not blocked. We can describe (See Table 3.1 and Expression (3.12)) this case as follows as a set of states in ρ_i^{Env}:

$$\{\mathcal{R}_i^{(\text{Env},j)} \mid j \in \{5, 9, 13\}\ \}.$$

  We translate both sub-cases in a corresponding diagonal matrix Λ_F as follows:

$$\Lambda_F \stackrel{\text{def}}{=} \text{diag}(0,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0)$$

  If we compare Λ_F with the diagonal matrix Λ_C which we defined earlier, then we can conclude that they are equal:

$$\Lambda_F = \Lambda_C.$$

3.6.3 The Transition Matrix $P^{B_i,\text{Cont}}$ for Transitions in ρ_i^{Cont} Without Transitions in ρ_i^{Env}

If we combine the results from the previous subsections we can construct the transition matrix P^{B_i} of the two-machine Markov model for buffer B_i excluding possible transitions in ρ_i^{Env}. We order the states in ρ_i^{Cont} in the following way:

$$(\ \mathcal{R}_i^{\text{Cont},0},\ \mathcal{R}_i^{\text{Cont},1},\ \mathcal{R}_i^{\text{Cont},2},\ ...,\ \mathcal{R}_i^{\text{Cont},N_i}\ ), \tag{3.14}$$

where inside $\mathcal{R}_i^{\text{Cont},j}$ we assume the natural order

$$\mathcal{R}_i^{\text{Cont},j} = \{\ (j,1),\ (j,2),\ (j,3),\ ...,\ (j,16)\}.$$

If we assume that transitions in $\rho_i^{Env_i}$ do not occur then we combine results from the previous paragraphs and construct the following transition matrix $P^{B_i,\text{Cont}}$:

$$P^{B_i,\text{Cont}} \stackrel{\text{def}}{=} \begin{pmatrix} \Lambda_A & \Lambda_C & & & & & & \\ \Lambda_B & \Lambda_D & \Lambda_C & & & & & \\ & \Lambda_E & \Lambda_D & \Lambda_C & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & \Lambda_E & \Lambda_D & \Lambda_C & & \\ & & & & \Lambda_E & \Lambda_D & \Lambda_C & \\ & & & & & \Lambda_E & \Lambda_D + \Lambda_E \end{pmatrix}. \tag{3.15}$$

3.6.4 The Transition Matrix $P^{B_i,\text{Env}}$ for Transitions in ρ_i^{Env} Without Transitions in ρ_i^{Cont}

If we use the transition matrix $P_{\rho_i^{\text{Env}}}$ which we constructed earlier, and use it as a building block, then we can construct the transition matrix P^{B_i} of the two-machine Markov model for buffer B_i excluding possible transitions in ρ_i^{Cont}. We order again the states in ρ_i^{Cont} as follows:

$$(\mathcal{R}_i^{\text{Cont},0},\ \mathcal{R}_i^{\text{Cont},1},\ \mathcal{R}_i^{\text{Cont},2},\ ...,\ \mathcal{R}_i^{\text{Cont},N_i}), \tag{3.16}$$

where inside $\mathcal{R}_i^{\text{Cont},j}$ we assume the natural order

$$\mathcal{R}_i^{\text{Cont},j} = \{\ (j,1),\ (j,2),\ (j,3),\ ...,\ (j,16)\ \}.$$

If we assume that transitions in ρ_i^{Cont} do not occur then we construct the following diagonal block structure in the following transition matrix $P^{B_i,\text{Env}}$:

$$P^{B_i,\text{Env}} \stackrel{\text{def}}{=} I_{N_i+1} \otimes P_{\rho_i^{\text{Env}}} = \begin{pmatrix} P_{\rho_i^{\text{Env}}} & & & & \\ & P_{\rho_i^{\text{Env}}} & & & \\ & & \ddots & & \\ & & & P_{\rho_i^{\text{Env}}} & \\ & & & & P_{\rho_i^{\text{Env}}} \end{pmatrix}. \tag{3.17}$$

3.6.5 The Transition Matrix P^{B_i} for Transitions in the State Set $\mathcal{R}^{B_i}$ with Transitions in ρ_i^{Env} and Transitions in ρ_i^{Cont} Combined

As we described earlier we will assume that within one time slot there will be *first* a transition in ρ_i^{Cont} and *next* a transition in ρ_i^{Env}. Since matrix multiplication is from the right to the left we get that the combined transition matrix P^{B_i} is a multiplication of the matrices $P^{B_i,\text{Env}}$ and $P^{B_i,\text{Cont}}$ as follows:

$$\begin{aligned} P^{B_i} &\stackrel{\text{def}}{=} P^{B_i,\text{Env}} \ .\ P^{B_i,\text{Cont}} \\ &= \left[I_{N_i+1} \otimes P_{\rho_i^{B_{i-1},M_{i-1}}} \otimes P_{\rho_i^{M_i,B_{i+1}}} \right] \cdot \begin{pmatrix} \Lambda_A & \Lambda_C & & & & & & \\ \Lambda_B & \Lambda_D & \Lambda_C & & & & & \\ & \Lambda_E & \Lambda_D & \Lambda_C & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & \Lambda_E & \Lambda_D & \Lambda_C & & \\ & & & & \Lambda_E & \Lambda_D & \Lambda_C & \\ & & & & & \Lambda_E & \Lambda_D + \Lambda_E \end{pmatrix}. \end{aligned} \tag{3.18}$$

where the following properties hold:

$$P_{\rho_i^{B_{i-1},M_{i-1}}},\ P_{\rho_i^{M_i,B_{i+1}}} \quad \in \quad \mathbb{R}^{4\times 4},$$

$$
\begin{aligned}
\Lambda_A,\ \Lambda_B &\in \mathbb{R}^{16\times 16}, \\
\Lambda_A + \Lambda_B &= I, \\
\Lambda_A\Lambda_B &= 0, \\
\Lambda_A^2 &= \Lambda_A, \\
\Lambda_B^2 &= \Lambda_B,
\end{aligned}
\qquad
\begin{aligned}
\Lambda_C,\ \Lambda_D,\ \Lambda_E &\in \mathbb{R}^{16\times 16}, \\
\Lambda_C + \Lambda_D + \Lambda_E &= I, \\
\Lambda_C\Lambda_D = \Lambda_C\Lambda_E = \Lambda_D\Lambda_E &= 0, \\
\Lambda_C^2 &= \Lambda_C, \\
\Lambda_D^2 &= \Lambda_D, \\
\Lambda_E^2 &= \Lambda_E,
\end{aligned}
\tag{3.19}
$$

and

$$
\begin{aligned}
\Lambda_A &= \mathrm{diag}(0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1), \\
\Lambda_B &= \mathrm{diag}(1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0), \\
\\
\Lambda_C &= \mathrm{diag}(0,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0), \\
\Lambda_D &= \mathrm{diag}(1,0,0,0,0,1,1,1,0,1,1,1,0,1,1,1), \\
\Lambda_E &= \mathrm{diag}(0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0).
\end{aligned}
\tag{3.20}
$$

3.7 The Stationary Distribution for the Two-Machine Markov Model at Buffer B_i

In this section we will examine the ways in which we can find the stationary distribution of the matrix P^{B_i} that we constructed in the previous section. In the following we will approach the problem of finding the stationary distribution in two different ways: first the analytic approach and then the numerical approach.

3.7.1 The Stationary Distribution Analytically

For the sake of simplicity of notation in this subsection the matrix M denotes the environment transition matrix $P_{\rho_i^{\mathrm{Env}}}$. We will assume that this transition matrix M is invertible. The problem for finding the stationary distribution can be formulated mathematically as follows: find the stationary distribution $\pi = (\pi_0'\ \pi_1'\ \pi_2'\ \ldots\quad \ldots\ \pi_N')'$ such that:

$$
\begin{pmatrix}
A & C & & & & & & \\
B & D & C & & & & & \\
 & E & D & C & & & & \\
 & & E & D & C & & & \\
 & & & \ddots & \ddots & \ddots & & \\
 & & & & E & D & C & \\
 & & & & & E & D & C \\
 & & & & & & E & D+E
\end{pmatrix}
\begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \\ \\ \\ \vdots \\ \pi_N \end{pmatrix}
=
\begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \\ \\ \\ \vdots \\ \pi_N \end{pmatrix},
\tag{3.21}
$$

where the following shorthand notations have been used:

$$
A \stackrel{\mathrm{def}}{=} M\Lambda_A,\ B \stackrel{\mathrm{def}}{=} M\Lambda_B,\ C \stackrel{\mathrm{def}}{=} M\Lambda_C,\ D \stackrel{\mathrm{def}}{=} M\Lambda_D,\ E \stackrel{\mathrm{def}}{=} M\Lambda_E. \tag{3.22}
$$

Similar problems, called "matrix geometric problems", were studied in [Neu81]. If the matrix M is non-singular then we can reformulate the main problem as follows:
find the stationary distribution $\pi = (\pi_0' \; \pi_1' \; \pi_2' \; \ldots \quad \ldots \; \pi_N')'$ such that:

$$\begin{pmatrix} \Lambda_A & \Lambda_C & & & & & & \\ \Lambda_B & \Lambda_D & \Lambda_C & & & & & \\ & \Lambda_E & \Lambda_D & \Lambda_C & & & & \\ & & \Lambda_E & \Lambda_D & \Lambda_C & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & \Lambda_E & \Lambda_D & \Lambda_C & \\ & & & & & \Lambda_E & \Lambda_D & \Lambda_C \\ & & & & & & \Lambda_E & \Lambda_D + \Lambda_E \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \\ \\ \\ \vdots \\ \pi_N \end{pmatrix} = \begin{pmatrix} M^{-1}\,\pi_0 \\ M^{-1}\,\pi_1 \\ \vdots \\ \\ \\ \\ \vdots \\ M^{-1}\,\pi_N \end{pmatrix}. \tag{3.23}$$

We have noticed earlier the following properties that are general properties of orthogonal projections on orthogonal subspaces:

- The Λ's are projections, and therefore idempotent:

$$\begin{aligned} \Lambda_A^2 &= \Lambda_A, \\ \Lambda_B^2 &= \Lambda_B, \\ \Lambda_C^2 &= \Lambda_C, \\ \Lambda_D^2 &= \Lambda_D, \\ \Lambda_E^2 &= \Lambda_E, \end{aligned} \tag{3.24}$$

- and the Λ's project on orthogonal subspaces:

$$\begin{aligned} \Lambda_A\,\Lambda_B &= \Lambda_B\,\Lambda_A = 0, \\ \Lambda_C\,\Lambda_D &= \Lambda_D\,\Lambda_C = 0, \\ \Lambda_D\,\Lambda_E &= \Lambda_E\,\Lambda_D = 0, \\ \Lambda_C\,\Lambda_E &= \Lambda_E\,\Lambda_C = 0. \end{aligned} \tag{3.25}$$

These properties give us the possibility to manipulate the equations in Expression (3.23). The equations in Expression (3.23) can be formulated as:

$$\Lambda_A\,\pi_0 + \Lambda_C\,\pi_1 = M^{-1}\,\pi_0, \tag{3.26}$$

$$\Lambda_B\,\pi_0 + \Lambda_D\,\pi_1 + \Lambda_C\,\pi_2 = M^{-1}\,\pi_1, \tag{3.27}$$

$$\Lambda_E\,\pi_{i-1} + \Lambda_D\,\pi_i + \Lambda_C\,\pi_{i+1} = M^{-1}\,\pi_i, \quad \text{for } 1 < i < N, \tag{3.28}$$

$$\Lambda_E\,\pi_{N-1} + (\Lambda_D + \Lambda_E)\,\pi_N = M^{-1}\,\pi_N. \tag{3.29}$$

Now we pre-multiply each of the equations in Expressions (3.26), (3.27), (3.28) and (3.29) by the matrices Λ_C, Λ_D and Λ_E and get using (3.24) and (3.25) that:

$$\begin{aligned}
\Lambda_C\,\Lambda_A\,\pi_0 + \Lambda_C\,\pi_1 &= \Lambda_C\,M^{-1}\,\pi_0,\\
\Lambda_D\,\Lambda_A\,\pi_0 &= \Lambda_D\,M^{-1}\,\pi_0,\\
\Lambda_E\,\Lambda_A\,\pi_0 &= \Lambda_E\,M^{-1}\,\pi_0,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\,\Lambda_B\,\pi_0 + \Lambda_C\,\pi_2 &= \Lambda_C\,M^{-1}\,\pi_1,\\
\Lambda_D\,\Lambda_B\,\pi_0 + \Lambda_D\,\pi_1 &= \Lambda_D\,M^{-1}\,\pi_1,\\
\Lambda_E\,\Lambda_B\,\pi_0 &= \Lambda_E\,M^{-1}\,\pi_1,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\,\pi_{i+1} &= \Lambda_C\,M^{-1}\,\pi_i, \quad \text{for } 1 < i < N,\\
\Lambda_D\,\pi_i &= \Lambda_D\,M^{-1}\,\pi_i, \quad \text{for } 1 < i < N,\\
\Lambda_E\,\pi_{i-1} &= \Lambda_E\,M^{-1}\,\pi_i, \quad \text{for } 1 < i < N,
\end{aligned}$$

$$\begin{aligned}
0 &= \Lambda_C\,M^{-1}\,\pi_N,\\
\Lambda_D\,\pi_N &= \Lambda_D\,M^{-1}\,\pi_N,\\
\Lambda_E\,(\pi_{N-1} + \pi_N) &= \Lambda_E\,M^{-1}\,\pi_N.
\end{aligned}$$

Next we *reorder* the previous set of equations and we obtain:

$$\begin{aligned}
\Lambda_D\,\Lambda_A\,\pi_0 &= \Lambda_D\,M^{-1}\,\pi_0,\\
\Lambda_E\,\Lambda_A\,\pi_0 &= \Lambda_E\,M^{-1}\,\pi_0,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\,\Lambda_A\,\pi_0 + \Lambda_C\,\pi_1 &= \Lambda_C\,M^{-1}\,\pi_0,\\
\Lambda_D\,\Lambda_B\,\pi_0 + \Lambda_D\,\pi_1 &= \Lambda_D\,M^{-1}\,\pi_1,\\
\Lambda_E\,\Lambda_B\,\pi_0 &= \Lambda_E\,M^{-1}\,\pi_1,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\,\Lambda_B\,\pi_0 + \Lambda_C\,\pi_2 &= \Lambda_C\,M^{-1}\,\pi_1,\\
\Lambda_D\,\pi_i &= \Lambda_D\,M^{-1}\,\pi_i, \quad \text{for } i = 2,\\
\Lambda_E\,\pi_{i-1} &= \Lambda_E\,M^{-1}\,\pi_i, \quad \text{for } i = 2,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\,\pi_{i+1} &= \Lambda_C\,M^{-1}\,\pi_i, \quad \text{for } 1 < i < N-1,\\
\Lambda_D\,\pi_i &= \Lambda_D\,M^{-1}\,\pi_i, \quad \text{for } 2 < i < N,\\
\Lambda_E\,\pi_{i-1} &= \Lambda_E\,M^{-1}\,\pi_i, \quad \text{for } 2 < i < N,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\,\pi_{i+1} &= \Lambda_C\,M^{-1}\,\pi_i, \quad \text{for } i = N-1,\\
\Lambda_D\,\pi_N &= \Lambda_D\,M^{-1}\,\pi_N,\\
\Lambda_E\,(\pi_{N-1} + \pi_N) &= \Lambda_E\,M^{-1}\,\pi_N,
\end{aligned}$$

$$0 \;=\; \Lambda_C\, M^{-1}\, \pi_N,$$

and if we adapt the counter i in some "blocks" of equations we get:

$$\begin{aligned}
\Lambda_D\, \Lambda_A\, \pi_0 &= \Lambda_D\, M^{-1}\, \pi_0, \\
\Lambda_E\, \Lambda_A\, \pi_0 &= \Lambda_E\, M^{-1}\, \pi_0,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\, \Lambda_A\, \pi_0 + \Lambda_C\, \pi_1 &= \Lambda_C\, M^{-1}\, \pi_0, \\
\Lambda_D\, \Lambda_B\, \pi_0 + \Lambda_D\, \pi_1 &= \Lambda_D\, M^{-1}\, \pi_1, \\
\Lambda_E\, \Lambda_B\, \pi_0 &= \Lambda_E\, M^{-1}\, \pi_1,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\, \Lambda_B\, \pi_0 + \Lambda_C\, \pi_2 &= \Lambda_C\, M^{-1}\, \pi_1, \\
\Lambda_D\, \pi_2 &= \Lambda_D\, M^{-1}\, \pi_2, \\
\Lambda_E\, \pi_1 &= \Lambda_E\, M^{-1}\, \pi_2,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\, \pi_i &= \Lambda_C\, M^{-1}\, \pi_{i-1}, \quad \text{for } 2 < i < N, \\
\Lambda_D\, \pi_i &= \Lambda_D\, M^{-1}\, \pi_i, \quad \text{for } 2 < i < N, \\
\Lambda_E\, \pi_{i-1} &= \Lambda_E\, M^{-1}\, \pi_i, \quad \text{for } 2 < i < N,
\end{aligned}$$

$$\begin{aligned}
\Lambda_C\, \pi_N &= \Lambda_C\, M^{-1}\, \pi_{N-1}, \\
\Lambda_D\, \pi_N &= \Lambda_D\, M^{-1}\, \pi_N, \\
\Lambda_E\, (\pi_{N-1} + \pi_N) &= \Lambda_E\, M^{-1}\, \pi_N,
\end{aligned}$$

$$0 \;=\; \Lambda_C\, M^{-1}\, \pi_N.$$

Next we add up the equations in the separate "blocks" and get:

$$\begin{aligned}
(\Lambda_D + \Lambda_E)\, \Lambda_A \pi_0 &= (\Lambda_D + \Lambda_E)\, M^{-1} \pi_0, \\
(\Lambda_C \Lambda_A + \Lambda_D \Lambda_B + \Lambda_E \Lambda_B)\pi_0 + (\Lambda_C + \Lambda_D)\pi_1 &= \Lambda_C M^{-1} \pi_0 + (\Lambda_D + \Lambda_E) M^{-1} \pi_1, \\
\Lambda_C \Lambda_B \pi_0 + \Lambda_E \pi_1 + (\Lambda_C + \Lambda_D)\pi_2 &= \Lambda_C M^{-1} \pi_1 + (\Lambda_D + \Lambda_E) M^{-1} \pi_2, \\
\text{for } 2 < i < N \quad , \Lambda_E \pi_{i-1} + (\Lambda_C + \Lambda_D)\pi_i &= \Lambda_C M^{-1} \pi_{i-1} + (\Lambda_D + \Lambda_E) M^{-1} \pi_i, \\
\Lambda_E \pi_{N-1} + \pi_N &= \Lambda_C M^{-1} \pi_{N-1} + (\Lambda_D + \Lambda_E) M^{-1} \pi_N, \\
0 &= \Lambda_C\, M^{-1}\, \pi_N.
\end{aligned} \tag{3.30}$$

From the last expressions we derive:

$$\begin{aligned}
(\Lambda_D + \Lambda_E)(\Lambda_A - M^{-1})\pi_0 &= 0, \\
\left[\Lambda_C + \Lambda_D - (\Lambda_D + \Lambda_E) M^{-1}\right] \pi_1 &= \left[\Lambda_C M^{-1} - (\Lambda_C \Lambda_A + \Lambda_D \Lambda_B + \Lambda_E \Lambda_B)\right] \pi_0, \\
\left[\Lambda_C + \Lambda_D - (\Lambda_D + \Lambda_E) M^{-1}\right] \pi_2 &= -\Lambda_C \Lambda_B \pi_0 + (\Lambda_C M^{-1} - \Lambda_E)\pi_1, \\
\left[\Lambda_C + \Lambda_D - (\Lambda_D + \Lambda_E) M^{-1}\right] \pi_i &= (\Lambda_C M^{-1} - \Lambda_E)\pi_{i-1}, \quad \text{for } 2 < i < N, \\
\left[I - (\Lambda_D + \Lambda_E) M^{-1}\right] \pi_N &= (\Lambda_C M^{-1} - \Lambda_E)\pi_{N-1} \\
\Lambda_C\, M^{-1}\, \pi_N &= 0.
\end{aligned} \tag{3.31}$$

If we now define

$$\begin{aligned} Y_0 &\stackrel{\text{def}}{=} (\Lambda_D + \Lambda_E)(\Lambda_A - M^{-1}), \\ Y_N &\stackrel{\text{def}}{=} \Lambda_C\, M^{-1}, \\ X &\stackrel{\text{def}}{=} \Lambda_C + \Lambda_D - (\Lambda_D + \Lambda_E)M^{-1}, \\ X_N &\stackrel{\text{def}}{=} I - (\Lambda_D + \Lambda_E)M^{-1}, \\ Z_0 &\stackrel{\text{def}}{=} \Lambda_C M^{-1} - (\Lambda_C\Lambda_A + \Lambda_D\Lambda_B + \Lambda_E\Lambda_B), \\ Z_1 &\stackrel{\text{def}}{=} -\Lambda_C\Lambda_B, \\ Z &\stackrel{\text{def}}{=} \Lambda_C M^{-1} - \Lambda_E, \end{aligned} \tag{3.32}$$

then we can reformulate the stationary distribution problem as follows:

$$\begin{aligned} Y_0\ \pi_0 &= 0, \\ X\ \pi_1 &= Z_0\ \pi_0, \\ X\ \pi_2 &= Z_1\ \pi_0 + Z\ \pi_1, \\ X\ \pi_i &= Z\ \pi_{i-1}, \quad \text{for } 2 < i < N, \\ X_N\ \pi_N &= Z\ \pi_{N-1}, \\ Y_N\ \pi_N &= 0. \end{aligned} \tag{3.33}$$

In matrix block form this is equal to:

$$\begin{pmatrix} Y_0 & & & & & & \\ -Z_0 & X & & & & & \\ -Z_1 & -Z & X & & & & \\ & & -Z & X & & & \\ & & & \ddots & \ddots & & \\ & & & & \ddots & X & \\ & & & & & -Z & X_N \\ & & & & & & Y_N \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_1 \\ \pi_2 \\ \vdots \\ \vdots \\ \pi_{N-2} \\ \pi_{N-1} \\ \pi_N \end{pmatrix} = 0. \tag{3.34}$$

The solution for four possible cases.

In the previous derivations we have shown that the problem can be reformulated as follows:

Find the *unique* vectors π_i for $0 \leq i \leq N$ such that

$$\begin{aligned} Y_0\ \pi_0 &= 0, \\ X\ \pi_1 &= Z_0\ \pi_0, \\ X\ \pi_2 &= Z_1\ \pi_0 + Z\ \pi_1, \\ X\ \pi_i &= Z\ \pi_{i-1}, \quad \text{for } 2 < i < N, \\ X_N\ \pi_N &= Z\ \pi_{N-1}, \\ Y_N\ \pi_N &= 0, \\ \sum_{i=0}^{N} \| \pi_i \|_1 &= 1. \end{aligned} \tag{3.35}$$

The solution to this kernel problem is simple in case the matrices X or X_N both are non-singular. The questions that arise naturally are the following two.

1. Do cases exist in which either X or X_N are singular?

2. If such cases exist, how does one solve the problem in these cases?

In the following we will show that all possible combinations of singular and non-singular X and X_N do exist. We also show how in each of these cases the *unique* solution can be found.

Case 1: X and X_N are both non-singular.

It is easy to find an example that proves that this case exists since this case consists of almost all random examples.
Example.

$$\begin{aligned} M &= \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{2} \end{pmatrix}, \quad \Lambda_A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ \Lambda_C &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{3.36}$$

leads to the following non-singular X, non-singular X_N and Z:

$$\begin{aligned} X &= \begin{pmatrix} 1 & 0 & 0 \\ -3 & 4 & -3 \\ 2 & -2 & 0 \end{pmatrix}, \quad X_N = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 4 & -3 \\ 2 & -2 & 1 \end{pmatrix}, \\ Z &= \begin{pmatrix} 0 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned} \tag{3.37}$$

□

If the matrices X and X_N are invertible then we can solve for the π_i, $i \in \{0, 1, 2, ..., N\}$ as

follows:

$$\begin{aligned}
Y_0\ \pi_0 &= 0, \\
\pi_1 &= X^{-1}\ Z_0\ \pi_0, \\
\pi_2 &= X^{-1}\ Z_1\ \pi_0 + X^{-1}\ Z\ \pi_1 \\
&= X^{-1}\ (Z_1 + Z\ X^{-1}\ Z_0)\ \pi_0, \\
\pi_i &= X^{-1}\ Z\ \pi_{i-1}, \quad \text{for } 2 < i < N, \\
&= [X^{-1}\ Z]^{i-2}\ X^{-1}\ (Z_1 + Z\ X^{-1}\ Z_0)\ \pi_0, \quad \text{for } 2 < i < N, \\
\pi_N &= X_N^{-1}\ Z\ \pi_{N-1} \\
&= X_N^{-1}\ Z\ [X^{-1}\ Z]^{N-3}\ X^{-1}\ (Z_1 + Z\ X^{-1}\ Z_0)\ \pi_0, \\
Y_N\ \pi_N &= Y_N\ X_N^{-1}\ Z\ [X^{-1}\ Z]^{N-3}\ X^{-1}\ (Z_1 + Z\ X^{-1}\ Z_0)\ \pi_0 \\
&= 0.
\end{aligned} \tag{3.38}$$

So the "solution" to the problem is first to find a π_0 such that:

$$\begin{aligned}
Y_0\ \pi_0 &= 0, \\
Y_N\ X_N^{-1}\ [Z\ X^{-1}]^{N-2}\ (Z_1 + Z\ X^{-1}\ Z_0)\ \pi_0 &= 0,
\end{aligned} \tag{3.39}$$

and then compute the other π_i for $0 < i \leq N$ by means of the following formula's:

$$\begin{aligned}
\pi_1 &= X^{-1}\ Z_0\ \pi_0, \\
\pi_2 &= X^{-1}\ Z_1\ \pi_0 + X^{-1}\ Z\ \pi_1, \\
\pi_i &= X^{-1}\ Z\ \pi_{i-1}, \quad \text{for } 2 < i < N, \\
\pi_N &= X_N^{-1}\ Z\ \pi_{N-1}.
\end{aligned} \tag{3.40}$$

Case 2: X and X_N are both singular.

An example that shows that this case exists is as follows.
Example.

$$\begin{aligned}
M &= \begin{pmatrix} \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{8} & 0 & \frac{3}{4} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}, \quad \Lambda_A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
\Lambda_B &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Lambda_C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
\Lambda_D &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
\end{aligned} \tag{3.41}$$

leads to the following singular X, singular X_N and Z:

$$
\begin{aligned}
X &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & 1 \\ 0 & 4 & 0 & -2 \end{pmatrix}, \quad X_N = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & 1 \\ 0 & 4 & 0 & -1 \end{pmatrix}, \\
Z &= \begin{pmatrix} -6 & 10 & 2 & -2 \\ 6 & -6 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
\end{aligned} \tag{3.42}
$$

□

The solution to the problem in this case is first to find the matrix L which has the maximal number of independent rows in the left nullspace of matrix X:

$$L\,X = 0. \tag{3.43}$$

If we pre-multiply the corresponding equations in Expression (3.35) by L then we obtain "extra" restrictions:

$$
\begin{aligned}
L\,X\,\pi_1 &= L\,Z_0\,\pi_0 &&= 0, \\
L\,X\,\pi_2 &= L\,Z_1\,\pi_0 + L\,Z\,\pi_1 &&= 0, \\
L\,X\,\pi_i &= L\,Z\,\pi_{i-1} &&= 0, \quad \text{for } 2 < i < N,
\end{aligned} \tag{3.44}
$$

If we combine the "extra" restrictions in (3.44) with the original equations in Expression (3.35) then we get:

$$
\begin{aligned}
\begin{pmatrix} Y_0 \\ L\,Z_0 \end{pmatrix} \pi_0 &= 0, \\
\begin{pmatrix} X \\ L\,Z \end{pmatrix} \pi_1 &= \begin{pmatrix} Z_0 \\ -L\,Z_1 \end{pmatrix} \pi_0, \\
\begin{pmatrix} X \\ L\,Z \end{pmatrix} \pi_2 &= \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_1 + \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} \pi_0, \\
\begin{pmatrix} X \\ L\,Z \end{pmatrix} \pi_i &= \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_{i-1}, \quad \text{for } 2 < i < N-1, \\
X\,\pi_{N-1} &= Z\,\pi_{N-2}, \\
\begin{pmatrix} X_N \\ Y_N \end{pmatrix} \pi_N &= \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_{N-1}.
\end{aligned} \tag{3.45}
$$

Notice that the matrix $\begin{pmatrix} X \\ L\ Z \end{pmatrix}$ has full column rank. This allows us to premultiply the corresponding equations with a left inverse of $\begin{pmatrix} X \\ L\ Z \end{pmatrix}$ denoted by K:

$$\begin{aligned} K &\stackrel{\text{def}}{=} (X'X + Z'L'LZ)^{-1} \ [X' \quad Z'L'], \\ K \begin{pmatrix} X \\ L\ Z \end{pmatrix} &= I. \end{aligned} \tag{3.46}$$

We obtain:

$$\begin{aligned} \begin{pmatrix} Y_0 \\ L\ Z_0 \end{pmatrix} \pi_0 &= 0, \\ \pi_1 &= K \begin{pmatrix} Z_0 \\ -L\ Z_1 \end{pmatrix} \pi_0, \\ \pi_2 &= K \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_1 + K \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} \pi_0, \\ \pi_i &= K \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_{i-1}, \quad \text{for } 2 < i < N-1. \\ X\ \pi_{N-1} &= Z\ \pi_{N-2}, \\ \begin{pmatrix} X_N \\ Y_N \end{pmatrix} \pi_N &= \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_{N-1}. \end{aligned} \tag{3.47}$$

The last equations give us the possibility to express π_{N-2} in terms of π_0:

$$\pi_{N-2} = \left\{ \left(K \begin{pmatrix} Z \\ 0 \end{pmatrix} \right)^{N-3} K \begin{pmatrix} Z_0 \\ -L\ Z_1 \end{pmatrix} + \left(K \begin{pmatrix} Z \\ 0 \end{pmatrix} \right)^{N-4} K \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} \right\} \pi_0.$$

If, for the sake of convenience of notation, we denote the matrix in the last expression by J:

$$J \stackrel{\text{def}}{=} \left(K \begin{pmatrix} Z \\ 0 \end{pmatrix} \right)^{N-3} K \begin{pmatrix} Z_0 \\ -L\ Z_1 \end{pmatrix} + \left(K \begin{pmatrix} Z \\ 0 \end{pmatrix} \right)^{N-4} K \begin{pmatrix} Z_1 \\ 0 \end{pmatrix}, \tag{3.48}$$

then we can express the solution to our problem as follows:

Find the vectors π_0, π_{N-1} and π_N such that the following equations hold:

$$\begin{pmatrix} Y_0 & & & \\ L\ Z_0 & & & \\ Z\ J & -X & & \\ & -Z & X_N & \\ & & Y_N & \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_{N-1} \\ \pi_N \end{pmatrix} = 0. \tag{3.49}$$

Then compute the other π_i for $i = 1, 2, 3, ..., N-2$ by means of the equations:

$$\begin{aligned} \pi_1 &= K \begin{pmatrix} Z_0 \\ -L\ Z_1 \end{pmatrix} \pi_0, \\ \pi_2 &= K \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_1 + K \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} \pi_0, \\ \pi_i &= K \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_{i-1}, \quad \text{for } 2 < i < N-1. \end{aligned} \tag{3.50}$$

Case 3: X non-singular and X_N singular.

An example that shows that this case exists is as follows.
Example.

$$\begin{aligned} M &= \begin{pmatrix} \frac{1}{8} & \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{8} & 0 & \frac{1}{2} & \frac{3}{4} \end{pmatrix}, \quad \Lambda_A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \Lambda_B &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Lambda_C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \Lambda_D &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{3.51}$$

which leads to the following non-singular X, singular X_N and Z:

$$\begin{aligned} X &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & -1 & 0 \\ -1 & -1 & 1 & -1 \end{pmatrix}, \quad X_N = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & -1 & 0 \\ -1 & -1 & 1 & 0 \end{pmatrix}, \\ Z &= \begin{pmatrix} -6 & 10 & -2 & 2 \\ 6 & -6 & 2 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \end{aligned} \tag{3.52}$$

□

The solution to the problem in this case is similar to the solution in the case that both matrices X and X_N are singular as we discussed in the previous Case 2. The difference

with the solution in the previous Case 2 is that, since matrix X is non-singular, here we matrix L is non-existent or in other words L is an "empty matrix". If we consider L to be an "empty matrix" we can translate the solution from Case 2 into the following solution:

We define

$$J \stackrel{\text{def}}{=} \left(X^{-1}\, Z\right)^{N-3}\, X^{-1}\, Z_0 + \left(X^{-1}\, Z\right)^{N-4}\, X^{-1}\, Z_1, \tag{3.53}$$

and we can express the solution to our problem as follows:

Find the vectors π_0, π_{N-1} and π_N such that the following equations hold:

$$\begin{pmatrix} Y_0 & & \\ Z\ J & -X & \\ & -Z & X_N \\ & & Y_N \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_{N-1} \\ \pi_N \end{pmatrix} = 0. \tag{3.54}$$

Then compute the other π_i for $i = 1, 2, 3, ..., N-2$ by means of the equations:

$$\begin{aligned} \pi_1 &= X^{-1}\, Z_0\ \pi_0, \\ \pi_2 &= X^{-1}\, Z\ \pi_1 + X^{-1}\, Z_1\ \pi_0, \\ \pi_i &= X^{-1}\, Z\ \pi_{i-1}, \quad \text{for } 2 < i < N-1. \end{aligned} \tag{3.55}$$

Case 4: X singular and X_N non-singular.

An example that shows that this case exists is as follows.
Example.

$$M = \begin{pmatrix} \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \\ \frac{1}{8} & 0 & \frac{3}{4} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad \Lambda_A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Lambda_B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Lambda_C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.56}$$

$$\Lambda_D = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_E = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

which leads to the following singular X, non-singular X_N and Z:

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 1 & 1 \\ -6 & -6 & 2 & 4 & 6 \\ -6 & -10 & 2 & 6 & 8 \end{pmatrix}, \quad X_N = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 1 & 1 \\ -6 & -6 & 2 & 5 & 6 \\ -6 & -10 & 2 & 6 & 9 \end{pmatrix},$$
$$Z = \begin{pmatrix} -6 & -6 & 2 & 6 & 6 \\ -6 & -10 & 2 & 6 & 10 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.57}$$

□

The solution to the problem in this case is similar to the solution in the case that both matrices X and X_N are singular as we discussed in Case 2 in this subsection:

First we define L as a matrix with a maximal number of independent rows in the left nullspace of matrix X:

$$L\,X = 0. \tag{3.58}$$

Then we define the matrices K and J as follows:

$$\begin{aligned} K &\stackrel{\text{def}}{=} (X'X + Z'L'LZ)^{-1} \; [X' \quad Z'L'], \\ J &\stackrel{\text{def}}{=} \left(K \begin{pmatrix} Z \\ 0 \end{pmatrix}\right)^{N-3} K \begin{pmatrix} Z_0 \\ -L\,Z_1 \end{pmatrix} + \left(K \begin{pmatrix} Z \\ 0 \end{pmatrix}\right)^{N-4} K \begin{pmatrix} Z_1 \\ 0 \end{pmatrix}, \end{aligned} \tag{3.59}$$

and then we can express the solution to our problem as follows:

Find the vectors π_0, π_{N-1} and π_N such that the following equations hold:

$$\begin{pmatrix} Y_0 & & & \\ L\,Z_0 & & & \\ Z\,J & -X & & \\ & -Z & X_N & \\ & & Y_N & \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_{N-1} \\ \pi_N \end{pmatrix} = 0. \tag{3.60}$$

Then compute the other π_i for $i = 1, 2, 3, ..., N-2$ by means of the equations:

$$\begin{aligned} \pi_1 &= K \begin{pmatrix} Z_0 \\ -L\ Z_1 \end{pmatrix} \pi_0, \\ \pi_2 &= K \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_1 + K \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} \pi_0, \\ \pi_i &= K \begin{pmatrix} Z \\ 0 \end{pmatrix} \pi_{i-1}, \quad \text{for } 2 < i < N-1. \end{aligned} \tag{3.61}$$

3.7.2 The Stationary Distribution Numerically

In this subsection we try to find the stationary distribution numerically. The stationary distribution, by definition, satisfies:

$$\begin{aligned} P^{B_i}\,\pi &= \pi\ , \\ \|\,\pi\,\|_1 &= 1, \end{aligned}$$

For the sake of simplicity of notation again in this subsection the matrix M denotes the environment transition matrix $P_{\rho_i^{\text{Env}}}$. We repeat the problem for finding the stationary distribution as follows: find the stationary distribution $\pi = (\pi_0'\ \pi_1'\ \pi_2'\ \ldots\quad \ldots\ \pi_N')'$ such that:

$$\begin{pmatrix} A & C & & & & & & \\ B & D & C & & & & & \\ & E & D & C & & & & \\ & & E & D & C & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & E & D & C & \\ & & & & & E & D & C \\ & & & & & & E & D+E \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \\ \\ \\ \vdots \\ \pi_N \end{pmatrix} = \begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \\ \\ \\ \vdots \\ \pi_N \end{pmatrix}, \tag{3.62}$$

where

$$A \stackrel{\text{def}}{=} M\Lambda_A\ ,\ B \stackrel{\text{def}}{=} M\Lambda_B\ ,\ C \stackrel{\text{def}}{=} M\Lambda_C\ ,\ D \stackrel{\text{def}}{=} M\Lambda_D\ ,\ E \stackrel{\text{def}}{=} M\Lambda_E.$$

It is known that the matrix P^{B_i} has exactly one eigenvalue equal to 1. We know also that P^{B_i}, as a well known property of stochastic matrices, has all (possibly complex) eigenvalues denoted by λ_j, $j \in \{1, 2, ..., 16(N_i+1)\}$, in the unit circle of the complex plane :

$$|\lambda_1| \leq |\lambda_2| \leq ... \leq |\lambda_{16(N_i+1)-1}| \leq \lambda_{16(N_i+1)} = 1,$$

$$\lambda_j \neq 1, \qquad \text{for } j \neq 16(N_i+1).$$

There also is exactly one eigenvector (stationary distribution) π corresponding to the eigenvalue 1. To solve for the stationary distribution we first subtract the matrix $(1+\delta)I, \delta \in \mathbb{R},\ 0 < \delta \ll 1$, from the original transition matrix P^{B_i}. The result is a *non-singular* tridiagonal matrix that has an eigenvalue equal to δ which is the smallest eigenvalue in the absolute sense. Next we solve the eigenvector that corresponds to this smallest eigenvalue δ by means of the inverse power method. The inverse power method implies for our case that we perform the the following iteration process:

$$\left(P^{B_i} - (1+\delta)I \right) \pi^{k+1} = \pi^k \ , \quad k = 0, 1, 2, .., \tag{3.63}$$

or, equivalently,

$$\begin{pmatrix} A-(1+\delta)I & C & & & & \\ B & D-(1+\delta)I & C & & & \\ & E & D-(1+\delta)I & C & & \\ & & \ddots & \ddots & \ddots & \\ & & & E & D-(1+\delta)I & C \\ & & & & E & D+E-(1+\delta)I \end{pmatrix} \begin{pmatrix} \pi_0^{k+1} \\ \pi_1^{k+1} \\ \vdots \\ \\ \\ \vdots \\ \pi_N^{k+1} \end{pmatrix} = \begin{pmatrix} \pi_0^{k} \\ \pi_1^{k} \\ \vdots \\ \\ \\ \vdots \\ \pi_N^{k} \end{pmatrix},$$

with an initial distribution $\pi_j^0 = \frac{1}{16(N_i+1)}(1,1,1,...,1)', \quad \forall j \in \{0,1,...,N_i\}$. Notice that the iteration procedure is *implicit*. We have to solve a linear set of $16(N_i+1)$ equations to solve for the next distribution π^{k+1} in each iteration step in (3.63). Luckily we can take advantage of the tridiagonal block structure of the matrix $P^{B_i} - (1+\delta)I$ in order to solve the equations: the iteration in Expression (3.63) can be executed with the help of the following blockwise LU-decomposition of $P^{B_i} - (1+\delta)I$:

$$P^{B_i} - (1+\delta)I = \begin{pmatrix} I & & & & & \\ Y_1 & I & & & & \\ & Y_2 & I & & & \\ & & Y_3 & I & & \\ & & & \ddots & \ddots & \\ & & & & Y_{N_i} & I \end{pmatrix} \begin{pmatrix} X_0 & C & & & & \\ & X_1 & C & & & \\ & & X_2 & C & & \\ & & & \ddots & \ddots & \\ & & & & X_{N_i-1} & C \\ & & & & & X_{N_i} \end{pmatrix},$$

where

$$\begin{aligned} X_0 &= A - (1-\delta)I \ , \\ Y_1 &= B\, X_0^{-1} \ , \end{aligned}$$

$$\left.\begin{aligned} X_k &= D - (1-\delta)I - Y_k C \ , \\ Y_{k+1} &= E\, X_k^{-1} \ , \end{aligned}\right\} \quad k = 1, 2, .., N_i - 1 \ ,$$

$$X_{N_i} = D + E - (1-\delta)I - Y_{N_i} C \ .$$

We implemented the procedure by means of the well known easy-to-use MATLAB software (short for *Mat*rix *Lab*oratory) by the company "The Mathworks Inc.". It turned out that our numerical procedure is quick and efficient. We used the value $\delta = 10^{-8}$ and found that four iteration steps from Expression (3.63) were more than enough to find a solution within machine accuracy.

3.8 Reduction of the Transition Matrix P^{B_i}

In Section 3.5 we described four main partitionings of which so far we used only the first two for the construction of the two-machine Markov model for buffer B_i. In this section we will use the third and fourth main partitionings for reduction of the two-machine Markov model for buffer B_i. Next to this reduction we will also introduce two functions and some notions that will make the descriptions in the following sections easier.

3.8.1 Reduction

Reduction to every possible partitioning of a Markov chain model is possible using the transition matrix and the stationary distribution as we have shown in Section 1.3.2. Both, the construction of the transition matrix P^{B_i} and the determination of its stationary distribution have been thoroughly described in the previous subsections. Therefore we now have all tools in hands for reduction. For reduction of Markov chains we will use the notation with a *tilde* and the corresponding partitioning as a *subscript* "$\tilde{(.)}_{\rho}$", introduced in Subsection 1.3.2. The two partitionings we focus on in this subsection correspond to partitionings $\rho_i^{M_{i-1},B_i}$ and $\rho_i^{B_i,M_i}$ in $\mathcal{P}(\mathcal{R}^{B_i})$. A picture of both partitionings is shown in Figure 3.4. Reduction to the transition matrices of the Markov chains shown in this figure, as a consequence of the notation introduced Subsection 1.3.2, can simply be denoted by $\tilde{P}^{B_i}_{\rho_i^{M_{i-1},B_i}}$ and $\tilde{P}^{B_i}_{\rho_i^{B_i,M_i}}$.

3.8.2 Construction, Determination of Stationary Distribution and Reduction Compressed in Two Functions

In the previous subsections we have followed the following procedure. We started with the two transition matrices $P_{\rho_i^{B_{i-1},M_{i-1}}}$ and $P_{\rho_i^{M_i,B_{i+1}}}$ corresponding to partitionings $\rho_i^{B_{i-1},M_{i-1}}$ and $\rho_i^{M_i,B_{i+1}}$ also shown in Figure 3.3. On the basis of these two transition matrices we constructed the transition matrix of the two-machine Markov chain model for buffer B_i. Then we described how the stationary distribution of the Markov chain model can be computed. Finally, in the previous subsection we described that by means of the transition matrix P^{B_i} and the corresponding stationary distribution π we can reduce the transition matrix P^{B_i} in two different ways, corresponding to two different partitionings $\rho_i^{M_{i-1},B_i}$ and $\rho_i^{B_i,M_i}$.

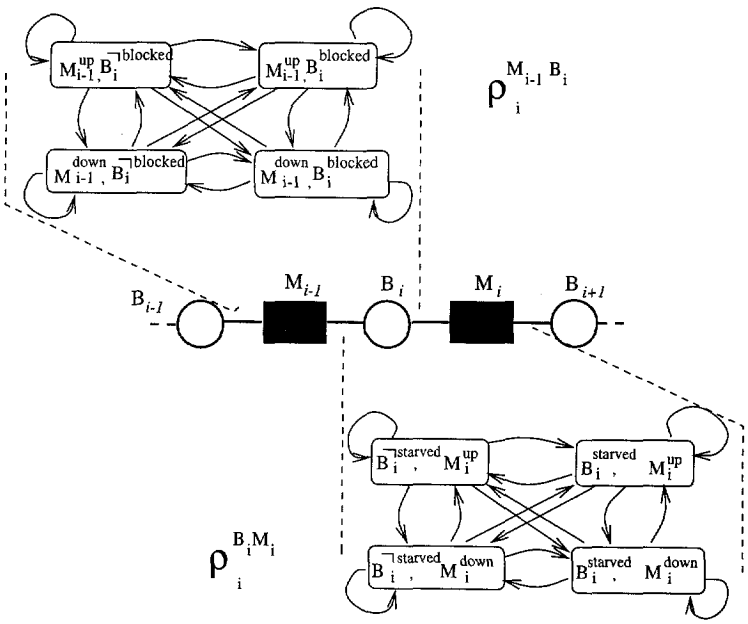

Figure 3.4: The two sub-chains for reduction to partitionings $\rho_i^{M_{i-1},B_i}$ and $\rho_i^{B_i,M_i}$.

Now we assume that instead of the "driving" Markov transition matrices $P_{\rho_i^{B_{i-1},M_{i-1}}}$ and $P_{\rho_i^{M_i,B_{i+1}}}$ we have arbitrary transition matrices X and Y. We can describe the transition matrix P^{B_i} as a function of the two "driving" Markov transition matrices $X \in \mathbb{R}^4 \times \mathbb{R}^4$ and $Y \in \mathbb{R}^4 \times \mathbb{R}^4$ similar to Expression (3.18):

$$P^{B_i}(X,Y) \stackrel{\text{def}}{=} [I_{N_i+1} \otimes X \otimes Y] \cdot \begin{pmatrix} \Lambda_A & \Lambda_C & & & & & & \\ \Lambda_B & \Lambda_D & \Lambda_C & & & & & \\ & \Lambda_E & \Lambda_D & \Lambda_C & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & \Lambda_E & \Lambda_D & \Lambda_C & & \\ & & & & \Lambda_E & \Lambda_D & \Lambda_C & \\ & & & & & \Lambda_E & \Lambda_D + \Lambda_E \end{pmatrix}, \tag{3.64}$$

where

$$\begin{aligned} \Lambda_A &= \text{diag}(0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1), \\ \Lambda_B &= \text{diag}(1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0), \\ \\ \Lambda_C &= \text{diag}(0,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0), \\ \Lambda_D &= \text{diag}(1,0,0,0,0,1,1,1,0,1,1,1,0,1,1,1), \\ \Lambda_E &= \text{diag}(0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0). \end{aligned}$$

The reduced transition matrices corresponding to partitionings $\rho_i^{M_{i-1},B_i}$ and $\rho_i^{B_i,M_i}$ as a function of the variable "driving" sub-chains X and Y can then be described concisely as $\tilde{P}^{B_i}_{\rho_i^{M_{i-1},B_i}}(X,Y) \in \mathbb{R}^4 \times \mathbb{R}^4$ and $\tilde{P}^{B_i}_{\rho_i^{B_i,M_i}}(X,Y) \in \mathbb{R}^4 \times \mathbb{R}^4$. In the following we will call $\tilde{P}^{B_i}_{\rho_i^{M_{i-1},B_i}}(X,Y)$ the *forward reduced Markov sub-chain* of the model at B_i and $\tilde{P}^{B_i}_{\rho_i^{B_i,M_i}}(X,Y)$ the *backward reduced Markov sub-chain* of the model at B_i. Transition matrices X and Y in this context will be called *"driving" Markov sub-chains* of the model at B_i.

3.9 Linkage of the Two-Machine Markov Models: the Forward and Backward Aggregation

Now we have fully described the construction, stationary distribution and reduction of the two-machine Markov model of buffer B_i. We did not yet describe the connections between two consecutive models: the model for buffer B_i and the model for buffer B_{i+1} for instance. In fact the way to link the models, using the tools we described earlier, is very

straightforward. Between two neighbouring buffer models we impose the following ties:

$$\begin{aligned} \tilde{P}^{B_{i+1}}_{\rho^{B_i,M_i}_{i+1}} &= \tilde{P}^{B_i}_{\rho^{B_i,M_i}_{i}}, \\ \tilde{P}^{B_i}_{\rho^{M_i,B_{i+1}}_{i}} &= \tilde{P}^{B_{i+1}}_{\rho^{M_i,B_{i+1}}_{i+1}}. \end{aligned} \tag{3.65}$$

Next we try to describe these ties in terms of the construction function in (3.18). The forward reduced Markov sub-chain from the model at B_i, denoted by $\tilde{P}^{B_i}_{\rho^{B_i,M_i}_{i}}(P_{\rho^{B_{i-1},M_{i-1}}_{i}}, P_{\rho^{M_i,B_{i+1}}_{i}})$, is equal to a "driving" Markov sub-chain for the model at B_{i+1}, denoted by $P_{\rho^{B_i,M_i}_{i+1}}$. Similarly, the backward reduced Markov sub-chain from the model at B_{i+1}, denoted by $\tilde{P}^{B_{i+1}}_{\rho^{M_i,B_{i+1}}_{i+1}}(P_{\rho^{B_i,M_i}_{i}}, P_{\rho^{M_{i+1},B_{i+2}}_{i}})$, is equal to a "driving" Markov sub-chain for the model at B_i, denoted by $P_{\rho^{M_i,B_{i+1}}_{i}}$.

In one expression we obtain the following equations equivalent with (3.65):

$$\begin{aligned} P_{\rho^{B_i,M_i}_{i+1}} &= \tilde{P}^{B_i}_{\rho^{B_i,M_i}_{i}}(P_{\rho^{B_{i-1},M_{i-1}}_{i}}, P_{\rho^{M_i,B_{i+1}}_{i}}), \\ P_{\rho^{M_i,B_{i+1}}_{i}} &= \tilde{P}^{B_{i+1}}_{\rho^{M_i,B_{i+1}}_{i+1}}(P_{\rho^{B_i,M_i}_{i+1}}, P_{\rho^{M_{i+1},B_{i+2}}_{i+1}}). \end{aligned} \tag{3.66}$$

Of course, in order to make sense on the level of the breakdown-repair models, we also have the following restrictions for all $i \in \{1, 2, ..., m\}$:

$$\left(\tilde{P}_{\rho^{B_i,M_i}_{i+1}}\right)_\mu = \left(\tilde{P}_{\rho^{M_i,B_{i+1}}_{i}}\right)_\mu = \begin{pmatrix} 1-\alpha_i & \beta_i \\ \alpha_i & 1-\beta_i \end{pmatrix}, \tag{3.67}$$

where

$$\mu = \{\ \{1,2\},\ \{3,4\}\ \} \in \mathcal{P}(\{1,2,3,4\}).$$

Next to the linking Equations (3.66) we also have boundary conditions for the backward reduced Markov sub-chain of the first and the forward reduced Markov sub-chain of the last machine. These sub-chains are fixed and are determined by the external process at the input buffer and the external process at the output buffer. We will assume in the following that starvation of the input buffer and blocking of the output buffer both do not occur. This can be expressed by means of the following boundary conditions:

$$\begin{aligned} P_{\rho^{B_1,M_1}_{2}} &= \begin{pmatrix} 1-\alpha_1 & 1-\alpha_1 & \beta_1 & \beta_1 \\ 0 & 0 & 0 & 0 \\ \alpha_1 & \alpha_1 & 1-\beta_1 & 1-\beta_1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ P_{\rho^{M_m,B_{m+1}}_{m}} &= \begin{pmatrix} 1-\alpha_m & 1-\alpha_m & \beta_m & \beta_m \\ 0 & 0 & 0 & 0 \\ \alpha_m & \alpha_m & 1-\beta_m & 1-\beta_m \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{3.68}$$

Notice that the structure of the combination of Expressions (3.66) and (3.68) is similar to the structure of Expressions (2.50) in the previous chapter about serial production lines with product failures only! The unknown matrices $P_{\rho_{i+1}^{B_i,M_i}}$ and $P_{\rho_i^{M_i,B_{i+1}}}$ in (3.66) and (3.68) correspond directly to the unknown scalar completion probabilities p_i^f and p_i^b in (2.50). The forward and backward aggregation iteration process to solve the equations in Expression (3.66) and (3.68) therefore will be chosen similar to the iteration process described in Section 2.4.8 to solve Equations (2.50). The forward and backward aggregation iteration process therefore can here be described similarly as follows:

- **Initialization**:

$$\begin{aligned} P_{\rho_{i+1}^{B_i,M_i}}(0) &= P_{\rho_i^{M_i,B_{i+1}}}(0) \\ &= \begin{pmatrix} 1-\alpha_i & 1-\alpha_i & \beta_i & \beta_i \\ 0 & 0 & 0 & 0 \\ \alpha_i & \alpha_i & 1-\beta_i & 1-\beta_i \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{for } i = 1, 2, .., m \ , \\ P_{\rho_2^{B_1,M_1}}(s) &= P_{\rho_2^{B_1,M_1}}(0) \ , \ \text{for } s = 1, 2, .. \\ P_{\rho_m^{M_m,B_{m+1}}}(s) &= P_{\rho_m^{M_m,B_{m+1}}}(0) \ , \ \text{for } s = 1, 2, .. \ . \end{aligned}$$

- **Forward aggregation**, for $i = 2, 3, .., m$:

$$P_{\rho_{i+1}^{B_i,M_i}}(s+1) = \tilde{P}^{B_i}_{\rho_i^{B_i,M_i}}(P_{\rho_i^{B_{i-1},M_{i-1}}}(s+1), P_{\rho_i^{M_i,B_{i+1}}}(s)). \tag{3.69}$$

- **Backward aggregation**, for $i = m, m-1, .., 2$.

$$P_{\rho_{i-1}^{M_{i-1},B_i}}(s+1) = \tilde{P}^{B_i}_{\rho_i^{M_{i-1},B_i}}(P_{\rho_i^{B_i,M_{i-1}}}(s+1), P_{\rho_i^{M_i,B_{i+1}}}(s+1)). \tag{3.70}$$

Note that although the alternate forward and backward iteration process is similar to the forward and backward process in Section 2.4.8, here we have not been able to prove the convergence of the algorithm or uniqueness of the solution.

3.10 Determination of the Average Production Rate R

Determination of the average production rate R is very straightforward. Again, as we have also seen in Subsection 2.4.3 in expressions (2.17) and (2.18), we can derive the average production rate of a two-machine model either from the average input rate or from the average output rate.

In order to determine the average input rate we examine in which states in the state set $\mathcal{R}^{B_i}$ a product part will enter the two-machine system model of buffer B_i. The conditions under which there will be input from the two-machine model at buffer B_i of a product part can be described as follows:

- The previous buffer B_{i-1} is not starved.
- The first machine in the model M_{i-1} is up.
- The buffer B_i is not blocked.

These conditions can be translated (See Table 3.1) in the following subset $\mathcal{R}_i^{\text{input}}$ of $\mathcal{R}^{B_i}$:

$$\mathcal{R}_i^{\text{input}} \stackrel{\text{def}}{=} \{(a,b) \in \mathcal{R}^{B_i} \mid a < N_i,\ b \in \{1,2,3,4\}\ \} \cup \{\ (N_i,1)\ \}. \tag{3.71}$$

The average input rate of buffer B_i can simply be obtained by adding up all elements of the stationary distribution that correspond with $\mathcal{R}_i^{\text{input}}$:

$$R_i^{\text{input}} \stackrel{\text{def}}{=} \sum_{(a,b)\in\mathcal{R}_i^{\text{input}}} \pi_{16a+b}, \tag{3.72}$$

where π is the stationary distribution such that:

$$P^{B_i}\ \pi\ =\ \pi.$$

In order to determine the average output rate we examine in which states in the state set $\mathcal{R}^{B_i}$ a product part will depart from the two-machine system model of buffer B_i. The conditions under which there will be output from the two-machine model at buffer B_i of a product part can be described as follows:

- The next buffer B_{i+1} is not blocked.
- The second machine in the model M_i is up.
- The buffer B_i is not starved.

These conditions can be translated (See Table 3.1) in the following subset $\mathcal{R}_i^{\text{output}}$ of $\mathcal{R}^{B_i}$:

$$\mathcal{R}_i^{\text{output}} \stackrel{\text{def}}{=} \{(a,b) \in \mathcal{R}^{B_i} \mid a > 0,\ b \in \{1,5,9,13\}\ \}. \tag{3.73}$$

The average output rate of buffer B_i can simply be obtained by adding up all elements of the stationary distribution that correspond with $\mathcal{R}_i^{\text{output}}$.

$$R_i^{\text{output}} \stackrel{\text{def}}{=} \sum_{(a,b)\in\mathcal{R}_i^{\text{output}}} \pi_{16a+b}, \tag{3.74}$$

where again π is the stationary distribution such that:

$$P^{B_i} \ \pi \ = \ \pi.$$

If the stationary distribution of the two-machine Markov model at buffer B_i has been computed correctly then the average *input* rate and the average *output* rate will be equal. They will be equal to *the* average production rate R_i of the model at buffer B_i:

$$R_i \ \stackrel{\text{def}}{=} \ R_i^{\text{input}} \ = \ R_i^{\text{output}}. \tag{3.75}$$

Notice however that the linkage equations in Expression (3.66) do *not* imply *conservation of flow*:

$$R_i \neq R_j, \qquad \text{for } i \neq j. \tag{3.76}$$

This is in contrast with the equations in Expression (2.50) for the production lines with product part failures only, as described in Chapter 2, where conservation of flow follows from the restrictions as shown in Theorem 29 in Appendix B. Since conservation of flow is not automatically satisfied here, we obtain several *different* estimates of the average production rate, as many estimates as there are two-machine production line models in the production line. If the estimates are good then the differences between the estimates from the production rates of the different two-machine models are very small. If the estimates are bad then the differences between the estimates from the production rates of the different two-machine models are big. In this way, the differences in flow are a measure for the accuracy of the two-machine production line models.

3.11 An Example

We apply the forward and backward iteration aggregation method described in this section to a simple serial production line with 4 machines and 5 buffers. A picture of this production line is shown in Figure 3.5. We can see in the picture that we are dealing with four machines $\mathcal{M} = \{M_1, M_2, M_3, M_4\}$ and five buffers $\mathcal{B} = \{B_1, B_2, B_3, B_4, B_5\}$. The capacities of buffers B_1, B_2, B_3, B_4 are ∞, 10, 20, 30, ∞ respectively. In the figure we also mention the parameters $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{11}$ and $\beta_1 = \frac{1}{2}$, $\beta_2 = \frac{1}{3}$, $\beta_3 = \frac{1}{4}$ and $\beta_4 = \frac{1}{5}$ which correspond to the breakdown and repair of machines M_1, M_2, M_3 and M_4. In the figure the breakdown and repair rates are written under the machines, first the breakdown and then the repair rate. The buffer capacities are written under the corresponding buffers. The idea behind the parameters in this example is to try to compensate for the effect of decreasing repair rates by increasing buffer capacities. One iteration of the iteration method consists of one forward aggregation as expressed in (3.69) and one backward aggregation as expressed in (3.70). We apply the iteration method to the production line with parameters shown in Figure 3.5 for 10 iteration steps ($s = 1, 2, .., 10$) and examine what happens.

First we examine the convergence of the iteration method and therefore we keep track of

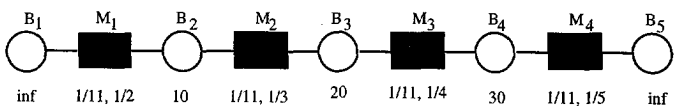

Figure 3.5: The production line in the example.

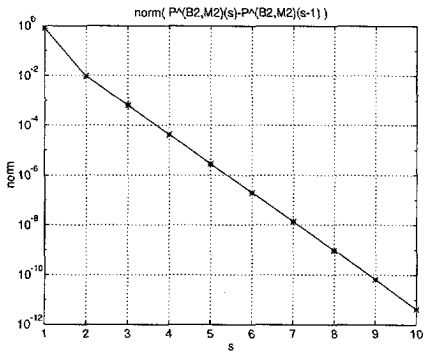

Figure 3.6: Logarithmic plot of $\| P_{\rho_3^{B_2,M_2}}(s) - P_{\rho_3^{B_2,M_2}}(s-1) \|$.

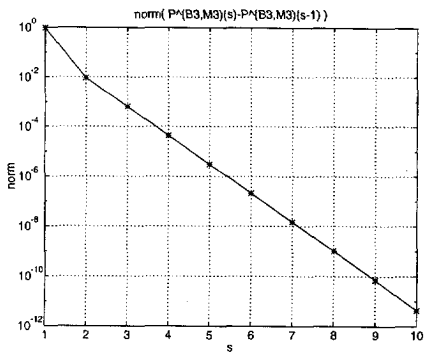

Figure 3.7: Logarithmic plot of $\| P_{\rho_4^{B_3,M_3}}(s) - P_{\rho_4^{B_3,M_3}}(s-1) \|$.

the subsequent differences of transition matrices along the iteration process. By taking the matrix norms of the subsequent differences $P_{\rho_{i+1}^{B_i,M_i}}(s) - P_{\rho_{i+1}^{B_i,M_i}}(s-1)$ and $P_{\rho_{i-1}^{M_{i-1},B_i}}(s) - P_{\rho_{i-1}^{M_{i-1},B_i}}(s-1)$ we are able to draw conclusions about the convergence of the iteration process for this case. For this simple serial production line we try to solve the following 6 different 4×4 Markov transition matrices: the three forward reduced Markov sub-chains $P_{\rho_3^{B_2,M_2}}$, $P_{\rho_4^{B_3,M_3}}$, $P_{\rho_5^{B_4,M_4}}$ and the three backward reduced Markov sub-chains $P_{\rho_3^{M_3,B_4}}$, $P_{\rho_2^{M_2,B_3}}$, $P_{\rho_1^{M_1,B_2}}$. Taking the norm of differences during the iteration process results in the plots on a logarithmic scale shown in Figures 3.6, 3.7, 3.8, 3.9, 3.10 and 3.11.

If we examine each single plot from Figures 3.6 up to 3.11 then we we conclude that the points of the differences in subsequent norms tend to lie on a straight line. Notice that if these points after some iteration steps lie on one monotonic descending line on the logarithmic scale then convergence of the iteration method is assured. Therefore the plots show that from the first 10 iteration steps we may conclude that convergence is attained for all matrices involved in the iteration method.

The approximations of the solutions to the combination of Expressions (3.66), (3.67) and

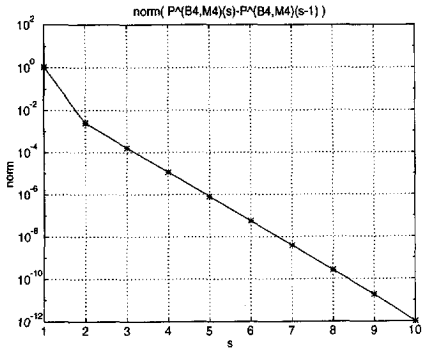

Figure 3.8: Logarithmic plot of $\| P_{\rho_5^{B_4,M_4}}(s) - P_{\rho_5^{B_4,M_4}}(s-1) \|$.

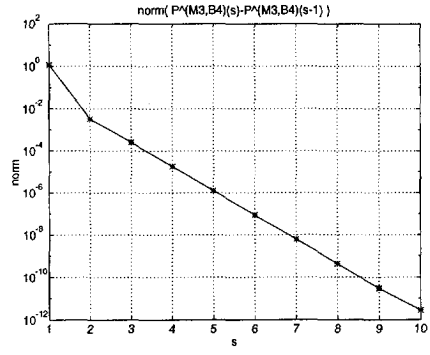

Figure 3.9: Logarithmic plot of $\| P_{\rho_3^{M_3,B_4}}(s) - P_{\rho_3^{M_3,B_4}}(s-1) \|$.

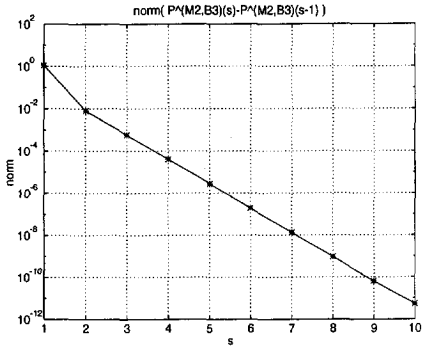

Figure 3.10: Logarithmic plot of $\| P_{\rho_2^{M_2,B_3}}(s) - P_{\rho_2^{M_2,B_3}}(s-1) \|$.

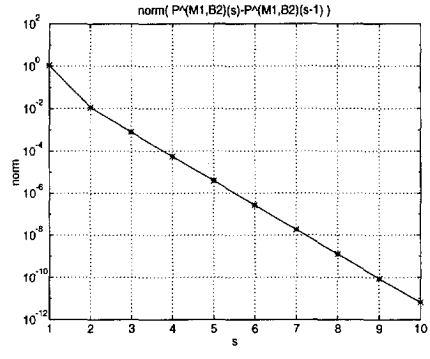

Figure 3.11: Logarithmic plot of $\| P_{\rho_1^{M_1,B_2}}(s) - P_{\rho_1^{M_1,B_2}}(s-1) \|$.

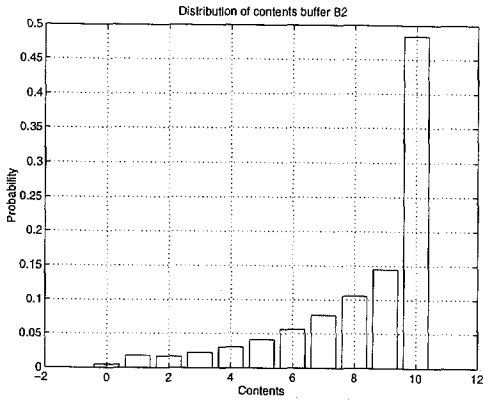

Figure 3.12: The *computed* stationary distribution of the contents of buffer B_2: $\tilde{\pi}_{\rho_2^{\text{Cont.}}}$.

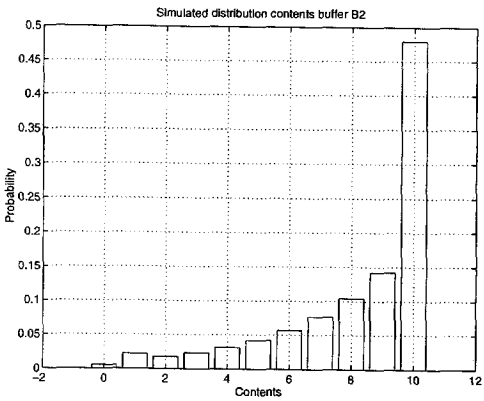

Figure 3.13: The *simulated* stationary distribution of the contents of buffer B_2.

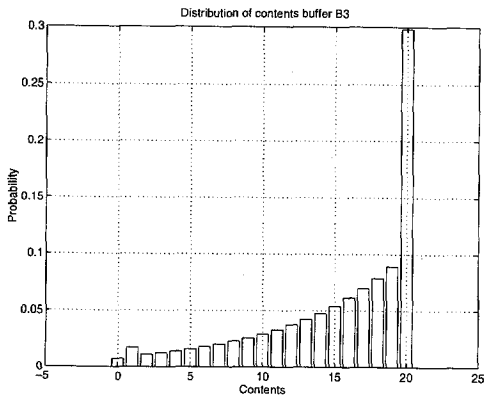

Figure 3.14: The *computed* stationary distribution of the contents of buffer B_3: $\tilde{\pi}_{\rho_3^{\text{Cont.}}}$.

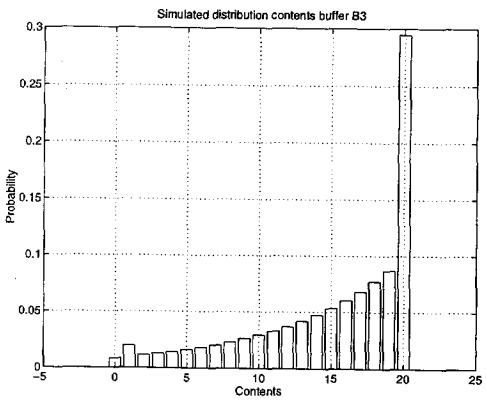

Figure 3.15: The *simulated* stationary distribution of the contents of buffer B_3.

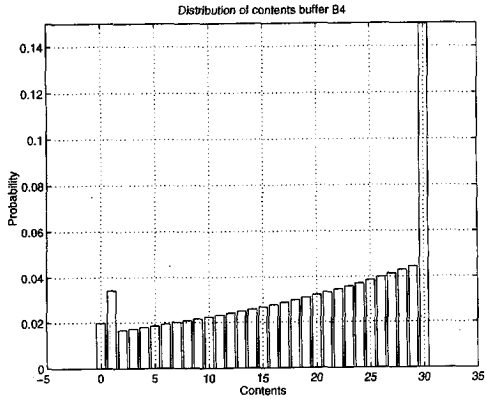

Figure 3.16: The *computed* stationary distribution of the contents of buffer B_4: $\tilde{\pi}_{\rho_4^{\text{Cont}}}$.

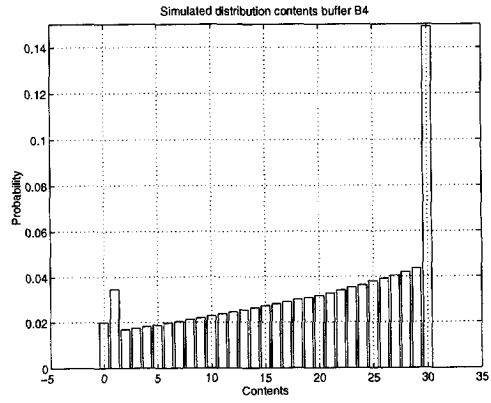

Figure 3.17: The *simulated* stationary distribution of the contents of buffer B_4.

(3.68) for this case after 10 steps of forward-backward iteration are:

$$P_{\rho_3^{B_2,M_2}}(10) = \begin{pmatrix} 0.9065 & 0.4545 & 0.3333 & 0.1667 \\ 0.0026 & 0.4545 & 0 & 0.1667 \\ 0.0907 & 0.0455 & 0.6667 & 0.3333 \\ 0.0003 & 0.0455 & 0 & 0.3333 \end{pmatrix},$$

$$P_{\rho_4^{B_3,M_3}}(10) = \begin{pmatrix} 0.9061 & 0.3041 & 0.2500 & 0.0836 \\ 0.0030 & 0.6050 & 0 & 0.1664 \\ 0.0906 & 0.0304 & 0.7500 & 0.2507 \\ 0.0003 & 0.0605 & 0 & 0.4993 \end{pmatrix},$$

$$P_{\rho_5^{B_4,M_4}}(10) = \begin{pmatrix} 0.9024 & 0.2275 & 0.2000 & 0.0500 \\ 0.0067 & 0.6815 & 0 & 0.1500 \\ 0.0902 & 0.0228 & 0.8000 & 0.2001 \\ 0.0007 & 0.0682 & 0 & 0.5999 \end{pmatrix},$$

$$P_{\rho_3^{M_3,B_4}}(10) = \begin{pmatrix} 0.8904 & 0.1818 & 0.2500 & 0.0500 \\ 0.0187 & 0.7273 & 0 & 0.2000 \\ 0.0890 & 0.0182 & 0.7500 & 0.1500 \\ 0.0019 & 0.0727 & 0 & 0.6000 \end{pmatrix},$$

$$P_{\rho_2^{M_2,B_3}}(10) = \begin{pmatrix} 0.8684 & 0.2074 & 0.3333 & 0.0756 \\ 0.0407 & 0.7017 & 0 & 0.2577 \\ 0.0868 & 0.0207 & 0.6667 & 0.1513 \\ 0.0041 & 0.0702 & 0 & 0.5154 \end{pmatrix},$$

$$P_{\rho_1^{M_1,B_2}}(10) = \begin{pmatrix} 0.8358 & 0.2450 & 0.5000 & 0.1337 \\ 0.0733 & 0.6641 & 0 & 0.3663 \\ 0.0836 & 0.0245 & 0.5000 & 0.1337 \\ 0.0073 & 0.0664 & 0 & 0.3663 \end{pmatrix}.$$

After the numerical computations of the backward and forward aggregation steps we are also able by means of simple manipulations of the different stationary distribution vectors to make plots of the stationary distribution of the buffer contents for each separate buffer. By means of these plots the usage and performance of each buffer can be examined. The computed distribution of the contents of the buffers B_2, B_3 and B_4 can be denoted by $\tilde{\pi}_{\rho_2^{\text{Cont}}}$, $\tilde{\pi}_{\rho_3^{\text{Cont}}}$ and $\tilde{\pi}_{\rho_4^{\text{Cont}}}$. They are shown in the plots in Figures 3.12, 3.14 and 3.16 respectively.

Of course, the solution to the restrictions of our decomposed production line does *not* give the exact stationary distribution of the original production line. In order to evaluate the accuracy of the previous solution we have also done some simulations of the original production line. Therefore we have made a program for simulation of our production line in C programming language. We use again the same program as we have used for simulations earlier for the production lines with product failures only, we only changed the parameters such that they suit for our example here. The most important tool for the simulations is the random generator used. For our simulations in C we used the *drand48* command which according to the manual pages generates pseudo-random numbers using a well-known linear congruential algorithm and 48-bit integer arithmetic. The pseudo-random number generator is initialized by means of a time dependent argument so as to generate an independent sequence each time it is invoked.

The results with respect to the stationary distribution of the buffer contents of buffers B_2, B_3 and B_4 are shown in Figures 3.13, 3.15 and 3.17. These figures are put next to the figures of the computed stationary distribution of the buffer contents for comparison. Other results concerning average production rate and average buffer contents of the 10 independent "runs" of simulation are shown in Table 3.2. The compound results are shown in Table 3.3.

□

Estimates of average production rate R by simulation										
	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
R	0.672420	0.669433	0.671803	0.672932	0.672168	0.671453	0.671952	0.671538	0.670532	0.670932
Estimates of the average buffer contents by simulation										
Buffer	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
B_2	8.2476	8.2551	8.2359	8.2393	8.2429	8.2556	8.2546	8.2409	8.2812	8.2473
B_3	14.9346	15.0348	15.0310	14.9436	15.1433	15.0402	15.1815	15.1242	15.1523	15.1616
B_4	18.3049	18.6510	18.4637	18.5147	18.5068	18.2870	18.4086	18.5036	18.6024	18.4881

Table 3.2: The results of 10^6 time steps simulations for the 4-machine-production line

Results for the average production rate R						
	mean value	std. deviation	decomp. model		abs. err.	rel. err.
R	0.6715163	0.0010095	B_2	0.6711962	-0.0003201	-0.0004767
			B_3	0.6714088	-0.0001075	-0.0001601
			B_4	0.6715286	0.0000123	0.0000183
Results for the average buffer contents						
Buffer	mean value	std. deviation	decomp. model		abs. err.	rel. err.
B_2	8.2500	0.0130	8.3138		0.0638	0.0077
B_3	15.0747	0.0904	15.2040		0.1293	0.0086
B_4	18.4731	0.1152	18.7134		0.2403	0.0130

Table 3.3: The compound results of the simulations for the 4-machine-production line

3.12 Conclusions

In this chapter we described a decomposition method for a discrete synchronous serial production line with machine failures only. The machines in these production lines are prone to breakdown and repair only. By means of the decomposition method presented here we are able to estimate the average production rate R of such production lines. The method constists of alternate forward and backward aggregation steps and has the same structure as the method described earlier in Chapter 2 for production lines with product failures only. In spite of some attempts, we have not been able to prove convergence of the algorithm and uniqueness of the solution. All examples, of which one is treated in this chapter, show that the algorithm does converge and that the method is surprisingly accurate. Accuracy can be examined by comparison with the results from simulations. We did not yet find examples in which the accuracy of the algorithm was poor, which of course does not mean that such examples do not exist.

The solutions from computations with the algorithm show that in general they do not satisfy exactly the well-known "conservation of flow" property of serial production lines with blocking. The lack of this property makes the method less elegant, but not necessarily less accurate. Computations of the iteration method have been done in the MATLAB software environment. In this environment the actual numerical computations such as inversion of 16×16 matrices is quick and efficient. Inefficient however is the memory management of the MATLAB software. That is why MATLAB spends too much of the total execution time in data management tasks. The computations for the example treated in the previous section for instance take some minutes of computation time on a HP 720 workstation. A tremendous speed up of the iteration method therefore could be obtained by translation of the method in a lower level programming language such as C. However, this was outside the scope of this research.

Due to the decomposition method there is a *linear* relation between the complexity of the algorithm and an increment in capacity of a buffer in the line. This property makes it possible to compute approximated versions of the stationary distribution even if the capacities of the buffers are large.

Chapter 4

Serial Production Lines with Machine and Product Failures Combined

In this chapter we will make an attempt to combine the model for product part failures only from Chapter 2 and the model for machine failures only from Chapter 3. This implies that we obtain a model in which machines have both features: the geometrically distributed service times and the breakdown and repair. Before we examine this kind of production line we will first describe how production lines with machine and product failures work.

4.1 Description of a Production Line with Machine and Product Failures

The production line with machine and product failures basically has all properties of the production line with machine failures only from the previous chapter, Chapter 3. The only difference with the model of machine failures only is the completion probability p_i, $0 < p_i < 1$ in the *up* state of machine M_i. Instead of a completion probability that switches between 1 and 0, as was the case in the model of machine failures only described in Section 3.1, here we assume that the completion probability of machine M_i switches between a machine dependent completion probability p_i for the *up* state and 0 for the *down* state. A picture of the Markov chain that governs the breakdown-repair model here is shown in Figure 4.1.

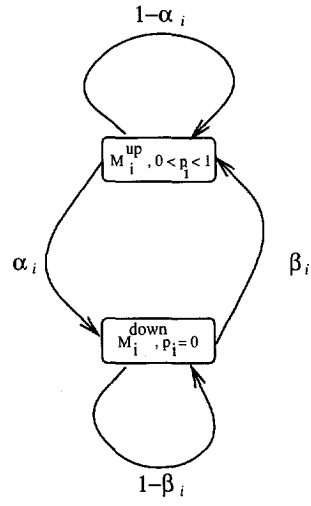

Figure 4.1: The breakdown-repair model of a machine M_i.

4.2 Adaptations to the Two-Machine Markov Chain Model for Production Lines with Machine Failures Only

Basically the structure of the two-machine Markov model for production lines with machine failures and product failures combined, is the same as the structure of the two-machine Markov model for production lines with machine failures only as we described in Chapter 3. That means that we assume that the two-machine Markov model for buffer B_i has again the same state set $\mathcal{R}^{B_i}$ as described in Table 3.1, and again we use the same partitionings in $\{ \rho_i^{B_{i-1},M_{i-1}}, \rho_i^{M_i,B_{i+1}}, \rho_i^{M_{i-1},B_i}, \rho_i^{\text{Env}}, \rho_i^{\text{Cont}_i} \} \subset \mathcal{P}(\mathcal{R}^{B_i})$ as defined in Expressions (3.7), (3.8), (3.9), (3.11) and (3.12).

One partitioning however, partitioning $\rho_i^{B_i,M_i}$, we did not mention in the previous paragraph. This is because in this partitioning $\rho_i^{B_i,M_i}$ we will introduce a change concerning states (0,1), (0,2), (0,3) and (0,3) in $\mathcal{R}^{B_i}$. In contrast with the definition in Expression (3.10), although buffer B_i is empty in these states, they will *not* be considered as starved states of buffer B_i. This is because in these states the fact that buffer B_i is empty will be of a "short term nature" since the previous machine is up and the previous buffer is not starved. For a buffer to be starved we will only take "long term arguments" into account. Therefore the new definition of $\rho_i^{B_i,M_i}$ will be:

$$\rho_i^{B_i,M_i} \stackrel{\text{def}}{=} (\mathcal{R}_i^{(B_i^{\neg\text{st}},M_i^{\text{up}})}, \mathcal{R}_i^{(B_i^{\text{st}},M_i^{\text{up}})}, \mathcal{R}_i^{(B_i^{\neg\text{st}},M_i^{\text{down}})}, \mathcal{R}_i^{(B_i^{\text{st}},M_i^{\text{down}})}) ,$$

where

$$\begin{aligned}
\mathcal{R}_i^{(B_i^{\neg st}, M_i^{up})} &= \{(a,b) \in \mathcal{R}^{B_i} | (a > 0 \wedge b \in \{1,2,5,6,9,10,13,14\}) \vee ... \\
&\qquad ...(a = 0 \wedge b \in \{1,2\})\}, \\
\mathcal{R}_i^{(B_i^{st}, M_i^{up})} &= \{(a,b) \in \mathcal{R}^{B_i} | a = 0 \wedge b \in \{5,6,9,10,13,14\}\}, \\
\mathcal{R}_i^{(B_i^{\neg st}, M_i^{down})} &= \{(a,b) \in \mathcal{R}^{B_i} | (a > 0 \wedge b \in \{3,4,7,8,11,12,15,16\}) \vee ... \\
&\qquad ...(a = 0 \wedge b \in \{3,4\})\}, \\
\mathcal{R}_i^{(B_i^{st}, M_i^{down})} &= \{(a,b) \in \mathcal{R}^{B_i} | a = 0 \wedge b \in \{7,8,11,12,15,16\}\}.
\end{aligned} \tag{4.1}$$

The rest of the description in Chapter 3 on production lines with machine failures only is also valid for our system here until the point where we describe the relation between the environment state ρ_i^{Env} and transitions in ρ_i^{Cont} in Subsection 3.6.2.

The relation between the environment state ρ_i^{Env} and transitions in ρ_i^{Cont} translated in the diagonal matrices Λ_A, Λ_B, Λ_C, Λ_D and Λ_E changes for the new model here. The new model will not have constant diagonal matrices Λ_A, Λ_B, Λ_C, Λ_D and Λ_E, but they will be dependent on two parameters p_1 and p_2 which denote the completion probabilities of the two machines involved in the model. The way in which the diagonal matrices depend on the completion probabilities p_1, $p_2 \in [0,1]$ and $q_1 = 1 - p_1$, $q_2 = 1 - p_2$ is as follows:

$$\begin{aligned}
\Lambda_A(p_1,p_2) &\stackrel{\text{def}}{=} \text{diag}(q_1,q_1,q_1,q_1,\ 1,1,1,1,\ 1,1,1,1,\ 1,1,1,1), \\
\Lambda_B(p_1,p_2) &\stackrel{\text{def}}{=} \text{diag}(p_1,p_1,p_1,p_1,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0), \\
\Lambda_C(p_1,p_2) &\stackrel{\text{def}}{=} \text{diag}(\ q_1p_2\ ,0,0,0,\ p_2,0,0,0,\ p_2,0,0,0,\ p_2,0,0,0), \\
\Lambda_D(p_1,p_2) &\stackrel{\text{def}}{=} \text{diag}(p_1p_2+q_1q_2,q_1,q_1,q_1,\ q_2,1,1,1,\ q_2,1,1,1,\ q_2,1,1,1), \\
\Lambda_E(p_1,p_2) &\stackrel{\text{def}}{=} \text{diag}(\ p_1q_2\ ,p_1,p_1,p_1,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0).
\end{aligned} \tag{4.2}$$

On the basis of these definitions, as an extension of Expression (3.15), we are now able to construct the new transition matrix $P^{B_i,\text{Cont}}$, for transitions in ρ_i^{Cont} without transitions in ρ_i^{Env}, as a part of the new two-machine Markov chain model for buffer B_i as follows:

$$P^{B_i,\text{Cont}}(p_{i-1}^f, p_i^b) \stackrel{\text{def}}{=}$$

$$\begin{pmatrix}
\Lambda_A(p_{i-1}^f,p_i^b) & \Lambda_C(p_{i-1}^f,p_i^b) & & & \\
\Lambda_B(p_{i-1}^f,p_i^b) & \Lambda_D(p_{i-1}^f,p_i^b) & \Lambda_C(p_{i-1}^f,p_i^b) & & \\
& \ddots & \ddots & \ddots & \\
& & \Lambda_E(p_{i-1}^f,p_i^b) & \Lambda_D(p_{i-1}^f,p_i^b) & \Lambda_C(p_{i-1}^f,p_i^b) \\
& & & \Lambda_E(p_{i-1}^f,p_i^b) & \Lambda_D(p_{i-1}^f,p_i^b) + \Lambda_E(p_{i-1}^f,p_i^b)
\end{pmatrix},$$

where p_{i-1}^f and p_i^b will be chosen such that they are in accordance with the forward and backward completion probabilities of "virtual machines" M_{i-1}^f and M_i^b as described in Expression (2.43) for lines with product failures only. How we will determine the forward and

backward completion probabilities p_{i-1}^f and p_i^b will be explained later in the next section, Section 4.3 in Expression (4.9).

Apart from the extension described in the previous paragraph, the new model described here is constructed in exactly the same way as we constructed the total transition matrix P^{B_i} for the two-machine Markov chain model for production lines with machine failures only. That means that we construct exactly the same transition matrix $P^{B_i,\mathrm{Env}}$, for transitions in ρ_i^{Env} without transitions in ρ_i^{Cont}, as in Expression (3.17):

$$\begin{aligned} P^{B_i,\mathrm{Env}} &\stackrel{\mathrm{def}}{=} I_{N_i+1} \otimes P_{\rho_i^{B_{i-1},M_{i-1}}} \otimes P_{\rho_i^{M_i,B_{i+1}}} \\ &= I_{N_i+1} \otimes P_{\rho_i^{\mathrm{Env}}} \\ &= \begin{pmatrix} P_{\rho_i^{\mathrm{Env}}} & & & & \\ & P_{\rho_i^{\mathrm{Env}}} & & & \\ & & \ddots & & \\ & & & P_{\rho_i^{\mathrm{Env}}} & \\ & & & & P_{\rho_i^{\mathrm{Env}}} \end{pmatrix}. \end{aligned} \tag{4.3}$$

The overall Markov transition matrix for the new extended model here then can be defined here in accordance with Expression (3.18).

$$\begin{aligned} P^{B_i} &\stackrel{\mathrm{def}}{=} P^{B_i,\mathrm{Env}} \cdot P^{B_i,\mathrm{Cont}}(p_{i-1}^f, p_i^b), \\ &= \left[I_{N_i+1} \otimes P_{\rho_i^{B_{i-1},M_{i-1}}} \otimes P_{\rho_i^{M_i,B_{i+1}}} \right] \begin{pmatrix} \Lambda_A(p_{i-1}^f, p_i^b) & \Lambda_C(p_{i-1}^f, p_i^b) & & \\ \ddots & \ddots & \ddots & \\ & \Lambda_E(p_{i-1}^f, p_i^b) & \Lambda_D(p_{i-1}^f, p_i^b) + \Lambda_E(p_{i-1}^f, p_i^b) \end{pmatrix}. \end{aligned} \tag{4.4}$$

Notice that the matrices Λ_A, Λ_B, Λ_C, Λ_D and Λ_E that we introduced in the earlier model from the previous chapter in Subsection 3.6.2 are consistent with the matrices that we obtain here if we take $p_{i-1}^f = 1$ and $p_i^b = 1$ in the new definitions:

$$\begin{aligned} \Lambda_A(1,1) &= \Lambda_A, \\ \Lambda_B(1,1) &= \Lambda_B, \\ \Lambda_C(1,1) &= \Lambda_C, \\ \Lambda_D(1,1) &= \Lambda_D, \\ \Lambda_E(1,1) &= \Lambda_E. \end{aligned}$$

This shows that the old two-machine Markov model for buffer B_i for machine failures only, as described in Chapter 3, is contained in the new two-machine Markov model for buffer B_i here.

On the other hand, if we consider a case in which for some positive values β_{i-1} and β_i

$$P^{B_{i-1},M_{i-1}}_{\rho_i} = \begin{pmatrix} 1 & 1 & 1-\beta_{i-1} & 1-\beta_{i-1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_{i-1} & \beta_{i-1} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad P^{M_i,B_{i+1}}_{\rho_i} = \begin{pmatrix} 1 & 1 & 1-\beta_i & 1-\beta_i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_i & \beta_i \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.5}$$

holds, then all states of $\mathcal{R}_i^{(\mathrm{Env},j)}$ of partitioning ρ_i^{Env} (See Expression (3.11)) except $\mathcal{R}_i^{(\mathrm{Env}_i,1)}$ will be transient. In such a case it is useful to consider the non-transient part of the two-machine Markov chain model. The transition probabilities of the non-transient part is determined by the (1,1) elements of the matrices $\Lambda_A(p_{i-1}^f, p_i^b)$, $\Lambda_B(p_{i-1}^f, p_i^b)$, $\Lambda_C(p_{i-1}^f, p_i^b)$, $\Lambda_D(p_{i-1}^f, p_i^b)$ and $\Lambda_E(p_{i-1}^f, p_i^b)$ which are q_{i-1}^f, p_{i-1}^f, $q_{i-1}^f p_i^b$, $p_{i-1}^f p_i^b + q_{i-1}^f q_i^b$ and $p_{i-1}^f q_i^b$ respectively. If we put these elements corresponding to the non-transient part of the Markov chain in a transition matrix we get the following structure for some values of p_{i-1}^f and p_i^b:

$$\begin{pmatrix} q_{i-1}^f & q_{i-1}^f p_i^b & 0 & 0 & \cdot & \cdot & 0 \\ p_{i-1}^f & p_{i-1}^f p_i^b + q_{i-1}^f q_i^b & q_{i-1}^f p_i^b & 0 & & & \cdot \\ 0 & p_{i-1}^f q_i^b & p_{i-1}^f p_i^b + q_{i-1}^f q_i^b & q_{i-1}^f p_i^b & & & \cdot \\ \cdot & & \ddots & \ddots & \ddots & & \cdot \\ \cdot & & & \ddots & \ddots & \ddots & \cdot \\ \cdot & & & & p_{i-1}^f q_i^b & p_{i-1}^f p_i^b + q_{i-1}^f q_i^b & q_{i-1}^f p_i^b \\ 0 & \cdot & \cdot & \cdot & 0 & p_{i-1}^f q_i^b & 1 - q_{i-1}^f p_i^b \end{pmatrix},$$

which corresponds directly to the two-machine production line model for production lines with product part failures only as described in Expression (2.11). We conclude that that for any buffer B_i the two machines production line model for product part failures only, as described in Chapter 2, is also contained in the combined two-machine Markov chain model that we describe now.

4.3 Linkage of the New Combined Model

For the linkage of the combined two-machine Markov chain models we first describe the combined two-machine Markov chain model as a function of its "driving" variables, in a similar way as did in Expression (3.18) for the two-machine Markov chain model for production lines with machine failures only. The difference with Expression (3.18) is that here we have to introduce two extra "driving variables" corresponding to the completion probabilities of both machines:

$$P^{B_i}(x, y, X, Y) \stackrel{\mathrm{def}}{=} \tag{4.6}$$

$$[I_{N_i+1} \otimes X \otimes Y] \begin{pmatrix} \Lambda_A(x,y) & \Lambda_C(x,y) & & & \\ \Lambda_B(x,y) & \Lambda_D(x,y) & \Lambda_C(x,y) & & \\ & \ddots & \ddots & \ddots & \\ & & \Lambda_E(x,y) & \Lambda_D(x,y)+\Lambda_E(x,y) & \end{pmatrix}.$$

For the link between the two-machine Markov models we can repeat the linking equations for the model for production lines with machine failures only, as described in Expression (3.66):

$$\begin{aligned} P_{\rho_{i+1}^{B_i,M_i}} &= \tilde{P}^{B_i}_{\rho_i^{B_i,M_i}}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}}), \\ P_{\rho_i^{M_i,B_{i+1}}} &= \tilde{P}^{B_{i+1}}_{\rho_{i+1}^{M_i,B_{i+1}}}(p_i^f, p_{i+1}^b, P_{\rho_{i+1}^{B_i,M_i}}, P_{\rho_i^{M_{i+1},B_{i+2}}}). \end{aligned} \tag{4.7}$$

Of course, in order to make sense on the level of the breakdown-repair models, we also have the following restrictions for all $i \in \{1,2,...,m\}$:

$$\left(\tilde{P}_{\rho_{i+1}^{B_i,M_i}}\right)_\mu = \left(\tilde{P}_{\rho_i^{M_i,B_{i+1}}}\right)_\mu = \begin{pmatrix} 1-\alpha_i & \beta_i \\ \alpha_i & 1-\beta_i \end{pmatrix}, \tag{4.8}$$

where

$$\mu = \{\ \{1,2\},\ \{3,4\}\ \} \in \mathcal{P}(\{1,2,3,4\}).$$

Next to these links we also have to make another link that determines the completion probabilities p_{i-1}^f and p_i^b of the "forward and backward aggregated machines" from the neighbouring two-machine models as an extension to the definitions in Expression (2.41).

It is difficult to extend Expression (2.41) because the notions "blocked" and "starved" have changed with the introduction of machine failures. We could argue that with the introduction of machine failures, we have two different kinds of "blocked" and "starved". We have a "*long term* blocked" and "*long term* starved" which means that a buffer is full or empty due to a *machine failure*. We also have a "*short term* blocked" and "*short term* starved" which means that a buffer is full or empty not due to a machine failure but due to a succession of *product part failures*.

For the production lines with product part failures only, as described in Chapter 2, we used the notions "blocked" and "starved" in a manner that we call "short term blocked" and "short term starved" now. For the production lines with machine failures only, as described in Chapter 3, we used the notions "blocked" and "starved" in a manner that we call "long term blocked" and "long term starved" now. It is important to realize the differences in meaning of the notions "blocked" and "starved" when we try to extend Expression (2.41) here.
The most useful extension that we could find is as follows:

$$\begin{aligned} p_i^b &\stackrel{\text{def}}{=} p_i.(1 - \mathbb{P}\text{r}(M_i \text{ is short term blocked})\ ), \\ p_i^f &\stackrel{\text{def}}{=} p_i.(1 - \mathbb{P}\text{r}(M_i \text{ is short term starved})\ ). \end{aligned}$$

or, equivalently,

$$\begin{array}{rcl} p_i^b & = & p_i(1 - \mathbb{P}\mathrm{r}(M_i \text{ is blocked} \mid M_i \text{ is up, not long term blocked or starved})), \\ p_i^f & = & p_i.(1 - \mathbb{P}\mathrm{r}(M_i \text{ is starved} \mid M_i \text{ is up, not long term blocked or starved})). \end{array} \tag{4.9}$$

In order to work this out more precisely we focus on the two-machine Markov chain model for a buffer B_i. All states in which machine M_{i-1} is up and M_{i-1} is neither long term blocked nor long term starved can be denoted by $\mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})} = \mathcal{R}_i^{(M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})} \cap \mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}})}$ (See also Table 3.1 and Expressions (3.7) and (3.9)):

$$\mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})} \stackrel{\mathrm{def}}{=} \{(a,b) \in \mathcal{R}^{B_i} | a < N_i \wedge b < 5\} \cup \{(N_i, 1)\}. \tag{4.10}$$

All states in which machine M_i is up and M_i is neither long term blocked nor long term starved can be denoted by $\mathcal{R}_i^{(B_i^{\neg\mathrm{st}}, M_i^{\mathrm{up}}, B_{i+1}^{\neg\mathrm{bl}})} = \mathcal{R}_i^{(B_i^{\neg\mathrm{st}}, M_i^{\mathrm{up}})} \cap \mathcal{R}_i^{(M_i^{\mathrm{up}}, B_{i+1}^{\neg\mathrm{bl}})}$ (See also Table 3.1 and Expressions (3.8) and (4.1)):

$$\mathcal{R}_i^{(B_i^{\neg\mathrm{st}}, M_i^{\mathrm{up}}, B_{i+1}^{\neg\mathrm{bl}})} \stackrel{\mathrm{def}}{=} \{(a,b) \in \mathcal{R}^{B_i} | (a > 0 \wedge b \in \{1,5,9,13\})\} \cup \{(0,1)\}. \tag{4.11}$$

By means of these definitions we are able to extend the expressions from Chapter 3 and find the corresponding expressions for the forward and backward completion probabilities p_i^f and p_i^b. In analogy with Expression (2.21), where conditional fractions take over the roles of π_N and π_0, we derive that

$$\mathbb{P}\mathrm{r}(M_{i-1} \text{ is blocked} \mid M_{i-1} \text{ is up, not long term blocked or starved}) =$$

$$q_i^b \cdot \frac{\mathbb{P}\mathrm{r}(B_i \text{ is in state } (N_i, 1) \in \mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})})}{\mathbb{P}\mathrm{r}(B_i \text{ is in a state from } \mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})})},$$

and

$$\mathbb{P}\mathrm{r}(M_i \text{ is starved} \mid M_i \text{ is up, not long term blocked or starved}) =$$

$$\frac{\mathbb{P}\mathrm{r}(B_i \text{ is in state } (0,1) \in \mathcal{R}_i^{(B_i^{\neg\mathrm{st}}, M_i^{\mathrm{up}}, B_{i+1}^{\neg\mathrm{bl}})})}{\mathbb{P}\mathrm{r}(B_i \text{ is in a state from } \mathcal{R}_i^{(B_i^{\neg\mathrm{st}}, M_i^{\mathrm{up}}, B_{i+1}^{\neg\mathrm{bl}})})}.$$

For the sake of convenience of notation we will define the following functions Ψ_i^{full} and Ψ_i^{empty}:

$$\Psi_i^{\mathrm{full}}(p_{i-1}^f, p_i^b, P_{\rho_i B_{i-1}, M_{i-1}}, P_{\rho_i M_i, B_{i+1}}) \stackrel{\mathrm{def}}{=}$$

$$\frac{\mathbb{P}\mathrm{r}(B_i \text{ is in state } (N_i, 1) \in \mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})})}{\mathbb{P}\mathrm{r}(B_i \text{ is in a state from } \mathcal{R}_i^{(B_{i-1}^{\neg\mathrm{st}}, M_{i-1}^{\mathrm{up}}, B_i^{\neg\mathrm{bl}})})}$$

$$= \frac{\pi_{16N_i+1}}{\sum_{(a,b)\in\mathcal{R}_i^{(B_{i-1}^{\neg st},M_{i-1}^{up},B_i^{\neg bl})}} \pi_{16a+b}}, \tag{4.12}$$

$$\Psi_i^{\text{empty}}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}}) \quad \stackrel{\text{def}}{=}$$

$$\frac{\mathbb{P}\text{r}(B_i \text{ is in state } (0,1) \in \mathcal{R}_i^{(B_i^{\neg st},M_i^{up},B_{i+1}^{\neg bl})})}{\mathbb{P}\text{r}(B_i \text{ is in a state from } \mathcal{R}_i^{(B_{i-1}^{\neg st},M_{i-1}^{up},B_i^{\neg bl})})}$$

$$= \frac{\pi_1}{\sum_{(a,b)\in\mathcal{R}_i^{(B_i^{\neg st},M_i^{up},B_{i+1}^{\neg bl})}} \pi_{16a+b}}, \tag{4.13}$$

where π is the stationary distribution vector of transition matrix $P^{B_i}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}})$:

$$P^{B_i}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}}) \cdot \pi = \pi.$$

The function Ψ_i^{full} could be considered as an "extension" of the function Ψ_N that we used in Chapter 2 and defined in Expression (2.15). We have used a similar symbol because the functions Ψ_i^{full} and Ψ_i^{empty} have similar interpretations as the function Ψ_N in Chapter 2.

As an extension of the formulas in Chapter 2 we can now derive for the model at B_i the following formulas in order to link the models on the level of the aggregated completion probabilities:

$$\begin{aligned} p_i^f &= p_i \cdot (1 - \Psi_i^{\text{empty}}), \\ p_{i-1}^b &= p_{i-1} \cdot (1 - q_i^b \cdot \Psi_i^{\text{full}}). \end{aligned} \tag{4.14}$$

(Here we have omitted arguments and simply wrote Ψ_i^{empty} and Ψ_i^{full})

To resume, arguments included, we can formulate the following linking equations:

$$\begin{aligned} P_{\rho_{i+1}^{B_i,M_i}} &= \tilde{P}_{\rho_i^{B_i,M_i}}^{B_i}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}}), \\ p_i^f &= p_i \cdot [1 - \Psi_i^{\text{empty}}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}})], \\ P_{\rho_i^{M_i,B_{i+1}}} &= \tilde{P}_{\rho_{i+1}^{M_i,B_{i+1}}}^{B_{i+1}}(p_i^f, p_{i+1}^b, P_{\rho_{i+1}^{B_i,M_i}}, P_{\rho_{i+1}^{M_{i+1},B_{i+2}}}), \\ p_i^b &= p_i \cdot [1 - q_{i+1}^b \cdot \Psi_{i+1}^{\text{full}}(p_i^f, p_{i+1}^b, P_{\rho_{i+1}^{B_i,M_i}}, P_{\rho_{i+1}^{M_{i+1},B_{i+2}}})]. \end{aligned} \tag{4.15}$$

with the following boundary values

$$
\begin{aligned}
P_{\rho_2^{B_1,M_1}} &= \begin{pmatrix} 1-\alpha_1 & 1-\alpha_1 & \beta_1 & \beta_1 \\ 0 & 0 & 0 & 0 \\ \alpha_1 & \alpha_1 & 1-\beta_1 & 1-\beta_1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
P_{\rho_m^{M_m,B_{m+1}}} &= \begin{pmatrix} 1-\alpha_m & 1-\alpha_m & \beta_m & \beta_m \\ 0 & 0 & 0 & 0 \\ \alpha_m & \alpha_m & 1-\beta_m & 1-\beta_m \\ 0 & 0 & 0 & 0 \end{pmatrix},
\end{aligned} \tag{4.16}
$$

$$
\begin{aligned}
p_1^f &= p_1 \\
p_m^b &= p_m,
\end{aligned}
$$

and restrictions

$$
\left(\tilde{P}_{\rho_{i+1}^{B_i,M_i}}\right)_\mu = \left(\tilde{P}_{\rho_i^{M_i,B_{i+1}}}\right)_\mu = \begin{pmatrix} 1-\alpha_i & \beta_i \\ \alpha_i & 1-\beta_i \end{pmatrix}, \tag{4.17}
$$

where

$$
\mu = \{\ \{1,2\},\ \{3,4\}\ \} \in \mathcal{P}(\{1,2,3,4\}).
$$

4.4 The Forward and Backward Aggregation

In order to solve the previous equations for matrices $P_{\rho_{i+1}^{B_i,M_i}}$ and $P_{\rho_i^{M_i,B_{i+1}}}$ and scalars p_i^f and p_i^b, such that they satisfy all Equations (4.15), (4.16) and (4.17), we will again use an iteration method that consists of successive forward and backward iteration steps. We use the variable s as the iteration counter as usually. The alternate forward and backward iteration method is as follows:

- **Initialization**:

$$
\begin{aligned}
P_{\rho_{i+1}^{B_i,M_i}}(0) &= P_{\rho_i^{M_i,B_{i+1}}}(0) \\
&= \begin{pmatrix} 1-\alpha_i & 1-\alpha_i & \beta_i & \beta_i \\ 0 & 0 & 0 & 0 \\ \alpha_i & \alpha_i & 1-\beta_i & 1-\beta_i \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{for } i = 1,2,..,m\ , \\
\left.\begin{aligned} P_{\rho_2^{B_1,M_1}}(s) &= P_{\rho_2^{B_1,M_1}}(0) \\ P_{\rho_m^{M_m,B_{m+1}}}(s) &= P_{\rho_m^{M_m,B_{m+1}}}(0) \end{aligned}\right\}, &\quad \text{for } s = 1,2,.. \\
p_i^f(0) = p_i^b(0) &= p_i\ , \quad \text{for } i = 1,2,..,m\ , \\
\left.\begin{aligned} p_1^f(s) &= p_1 \\ p_m^b(s) &= p_m \end{aligned}\right\}, &\quad \text{for } s = 1,2,..
\end{aligned}
$$

- **Forward aggregation**, for $i = 2, 3, .., m$:

$$\begin{aligned} P_{\rho_{i+1}^{B_i,M_i}}(s+1) \quad &= \\ & \tilde{P}^{B_i}_{\rho_i^{B_i,M_i}}(p_{i-1}^f(s+1), p_i^b(s), P_{\rho_i^{B_{i-1},M_{i-1}}}(s+1), P_{\rho_i^{M_i,B_{i+1}}}(s)), \\ p_i^f(s+1) \quad &= \\ & p_i \cdot [1 - \Psi_i^{\text{empty}}(p_{i-1}^f(s+1), p_i^b(s), P_{\rho_i^{B_{i-1},M_{i-1}}}(s+1), P_{\rho_i^{M_i,B_{i+1}}}(s))]. \end{aligned} \tag{4.18}$$

- **Backward aggregation**, for $i = m, m-1, .., 2$.

$$\begin{aligned} P_{\rho_{i-1}^{M_{i-1},B_i}}(s+1) \quad &= \\ & \tilde{P}^{B_i}_{\rho_i^{M_{i-1},B_i}}(p_{i-1}^f(s+1), p_i^b(s+1), P_{\rho_i^{B_{i-1},M_{i-1}}}(s+1), P_{\rho_i^{M_i,B_{i+1}}}(s+1)), \\ p_{i-1}^b(s+1) \quad &= \\ & p_{i-1} \cdot [1 - q_i^b(s+1) \cdot ... \\ & ...\Psi_i^{\text{full}}(p_{i-1}^f(s+1), p_i^b(s+1), P_{\rho_i^{B_{i-1},M_{i-1}}}(s+1), P_{\rho_i^{M_i,B_{i+1}}}(s+1))]. \end{aligned} \tag{4.19}$$

4.5 Determination of the Average Production Rate R

Determination of the average production rate R is straightforward. Again, as we have also seen in Subsection 2.4.3 in Expressions (2.17) and (2.18) as well as in Section 3.10 in Expressions (3.72) and (3.74), we can derive the average production rate of a two-machine model either from the average input rate or from the average output rate.

4.5.1 The Average Input Rate

In order to determine the average input rate we examine under which conditions a product part will enter the two-machine system of buffer B_i. The conditions under which there will be *input* from the two-machine model at buffer B_i of a product part can be split in two different cases as follows:

1. The previous buffer B_{i-1} is not long term starved, the first machine in the model M_{i-1} is up, completes a product part with probability p_{i-1}^f and buffer B_i contains less than N_i product parts.

2. The previous buffer B_{i-1} is not long term starved, the first machine in the model M_{i-1} is up, it completes a product part with probability p_{i-1}^f, buffer B_i contains N_i product parts and the second machine in the model M_i is up, completes a product part as well with probability p_i^b and the last buffer B_{i+1} is not long term blocked.

The probability that the first case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 1}) = p_{i-1}^f \cdot (\sum_{\substack{(a,b) \in \mathcal{R}^{B_i} \\ a < N_i \\ b < 5}} \pi_{16a+b}), \tag{4.20}$$

the probability that second case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 2}) = p_{i-1}^f \cdot \pi_{16N_i+1} \cdot p_i^b, \tag{4.21}$$

where π is the stationary distribution vector such that:

$$P^{B_i}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}}) \cdot \pi = \pi.$$

In total, with the same distribution vector π, we can now define the average input rate of the two-machine model as follows as the sum of both cases:

$$R_i^{\text{input}} \stackrel{\text{def}}{=} p_{i-1}^f \cdot (p_i^b \cdot \pi_{16N_i+1} + \sum_{\substack{(a,b) \in \mathcal{R}^{B_i} \\ a < N_i \\ b < 5}} \pi_{16a+b}). \tag{4.22}$$

4.5.2 The Average Output Rate

The conditions under which there will be *output* from the two-machine model at buffer B_i of a product part can be described in one case as follows:

1. Buffer B_i contains at least one product part, the second machine in the model M_i is up, completes a product part with probability p_i^b and the next buffer B_{i+1} is not long term blocked.

If we translate this in a formula we obtain:

$$R_i^{\text{output}} \stackrel{\text{def}}{=} p_i^b \cdot (\sum_{\substack{(a,b) \in \mathcal{R}^{B_i} \\ a > 0 \\ b \in \{1,5,9,13\}}} \pi_{16a+b}), \tag{4.23}$$

where again π is the stationary distribution vector such that:

$$P^{B_i}(p_{i-1}^f, p_i^b, P_{\rho_i^{B_{i-1},M_{i-1}}}, P_{\rho_i^{M_i,B_{i+1}}}) \cdot \pi = \pi.$$

4.5.3 Conservation of Flow

If the stationary distribution of the two-machine Markov model at buffer B_i has been computed correctly then the average *input* rate R_i^{input} and the average *output* rate R_i^{output} will be equal:

$$R_i^{\text{input}} = R_i^{\text{output}}. \tag{4.24}$$

They will be equal to the average production rate of the model at buffer B_i.

Notice however that the linkage equations in Expression (4.15) do **not** imply *conservation of flow*. This is in contrast with the equations in Expression (2.50) where conservation of flow follows from the restrictions as shown in Theorem 3 in part 29. Since conservation of flow is not automatically satisfied, we obtain several *different* estimates of the average production rate, as many estimates as there are two-machine production line models in the production line. If the estimates are good then the differences between the estimates from the production rates of the different two-machine models will be very small. If the estimates are bad then the differences between the estimates from the production rates of the different two-machine models will be big. In this way, the differences in flow are a measure for the accuracy of the two-machine production line models.

4.6 An Example

Again we will apply the forward and backward iteration aggregation method described in the previous section to a simple serial production line with 4 machines and 5 buffers. The production line in this example is similar to the production of the example in the previous chapter, except that the machines here have completion probabilities as well. A picture of this production line is shown in Figure 4.2. We can see in the picture that we are again dealing with four machines $\mathcal{M} = \{M_1, M_2, M_3, M_4\}$ and five buffers $\mathcal{B} = \{B_1, B_2, B_3, B_4, B_5\}$. Again the capacities of buffers B_1, B_2, B_3, B_4 are ∞, 10, 20, 30, ∞ respectively. Also the parameters for breakdown and repair rates of machines M_1, M_2, M_3 and M_4 are the same, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{11}$ and $\beta_1 = \frac{1}{2}$, $\beta_2 = \frac{1}{3}$, $\beta_3 = \frac{1}{4}$ and $\beta_4 = \frac{1}{5}$. In the figure the breakdown and repair rates are written under the machines, first the breakdown and then the repair rate. The buffer capacities are written under the corresponding buffers. The completion probabilities introduced for machines M_1, M_2, M_3 and M_4 are $p_1 = 0.40625$, $p_2 = 0.50000$, $p_3 = 0.50000$ and $p_4 = 0.50000$ respectively.

The first idea behind the introduction of completion probabilities is to see what happens if completion probabilities are in the middle between 0 and 1. It turned out that the case of taking all completion probabilities $p_i = 0.5$ was not interesting enough because the last machine was too "simple" a *bottleneck* because of its smallest repair rate. Therefore we introduced a second bottleneck at the first machine M_1 by means of a smaller completion probability $p_1 = 0.40625$. In this way we created two different bottlenecks in the same line:

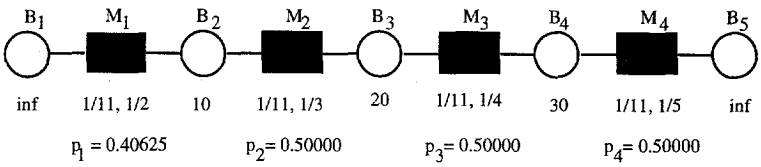

Figure 4.2: The production line in the example.

a bottleneck for the completion probability at the first machine M_1 and a bottleneck for the repair rates at the last machine M_4. The two bottlenecks are chosen such that both the first and last machine, considered in isolation, would produce exactly the same average amount of products per time slot.

With the introduction of the second bottleneck the number of iterations needed for accurate results increased. It turned out that for 8-digits accurate results here we needed at least 19 subsequent forward and backward iteration steps (s=0,1,..,19). First we examine the convergence of the iteration method and therefore we keep track of the subsequent differences of transition matrices along the iteration process. By taking the matrix norms of the subsequent differences $P_{\rho_{i+1}^{B_i,M_i}}(s)-P_{\rho_{i+1}^{B_i,M_i}}(s-1)$ for $i=2,3,4$ and $P_{\rho_i^{M_i,B_{i+1}}}(s)-P_{\rho_i^{M_i,B_{i+1}}}(s-1)$ for $i=1,2,3$ we are able to draw conclusions about the convergence of the iteration process for this case. For this simple serial production line we try to solve the following 6 different 4×4 Markov transition matrices: the tree forward reduced Markov sub-chains $P_{\rho_3^{B_2,M_2}}$, $P_{\rho_4^{B_3,M_3}}$, $P_{\rho_5^{B_4,M_4}}$ and the three backward reduced Markov sub-chains $P_{\rho_3^{M_3,B_4}}$, $P_{\rho_2^{M_2,B_3}}$, $P_{\rho_1^{M_1,B_2}}$. New in our model are the "aggregated" forward and backward completion probabilities that we also try to solve in our iteration process: the three forward aggregated completion probabilities p_2^f, p_3^f and p_4^f and the three backward aggregated completion probabilities p_3^b, p_2^b and p_1^b. Taking the norm of subsequent differences during the iteration process results in the plots on a logarithmic scale shown in Figures 4.3, 4.4, 4.5, 4.6. The solutions for all forward and backward reduced Markov chains, as well as the solutions for all aggregated forward and backward completion probabilities after 19 iteration steps are listed below.

$$P_{\rho_3^{B_2,M_2}}(19) = \begin{pmatrix} 0.8959 & 0.4545 & 0.3322 & 0.1667 \\ 0.0132 & 0.4545 & 0.0012 & 0.1667 \\ 0.0896 & 0.0455 & 0.6643 & 0.3333 \\ 0.0013 & 0.0455 & 0.0024 & 0.3333 \end{pmatrix}, \quad p_2^f(19) = 0.4477,$$

$$P_{\rho_4^{B_3,M_3}}(19) = \begin{pmatrix} 0.9011 & 0.3071 & 0.2497 & 0.0843 \\ 0.0080 & 0.6020 & 0.0003 & 0.1657 \\ 0.0901 & 0.0307 & 0.7490 & 0.2529 \\ 0.0008 & 0.0602 & 0.0010 & 0.4971 \end{pmatrix}, \quad p_3^f(19) = 0.4790,$$

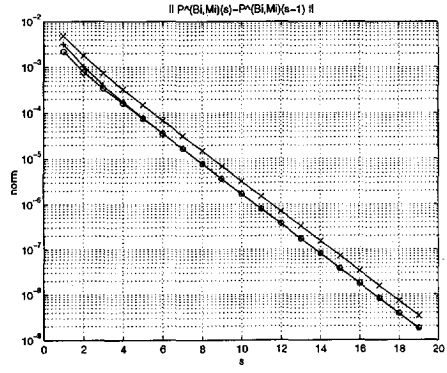

Figure 4.3: $\| P_{\rho_{i+1}^{B_i,M_i}}(s) - P_{\rho_{i+1}^{B_i,M_i}}(s-1) \|$ for i=2,3,4 we used "o","x","+" respectively.

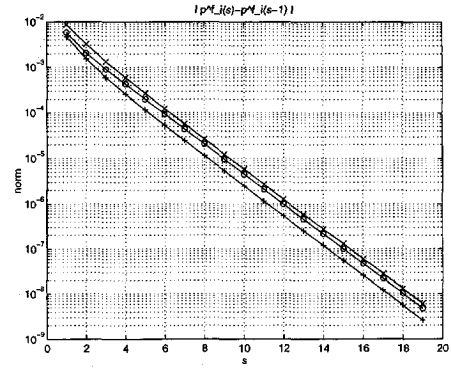

Figure 4.4: $\mid p_i^f(s) - p_i^f(s-1) \mid$ for i=2,3,4 we used "o","x","+" respectively.

$$P_{\rho_5^{B_4,M_4}}(19) = \begin{pmatrix} 0.9038 & 0.2274 & 0.1999 & 0.0500 \\ 0.0053 & 0.6817 & 0.0001 & 0.1500 \\ 0.0904 & 0.0227 & 0.7995 & 0.1999 \\ 0.0005 & 0.0682 & 0.0005 & 0.6001 \end{pmatrix}, \quad p_4^f(19) = 0.4889.$$

$$P_{\rho_3^{M_3,B_4}}(19) = \begin{pmatrix} 0.9033 & 0.1818 & 0.2498 & 0.0500 \\ 0.0058 & 0.7273 & 0.0002 & 0.2000 \\ 0.0903 & 0.0182 & 0.7495 & 0.1500 \\ 0.0006 & 0.0727 & 0.0005 & 0.6000 \end{pmatrix}, \quad p_3^b(19) = 0.4927,$$

$$P_{\rho_2^{M_2,B_3}}(19) = \begin{pmatrix} 0.9032 & 0.2184 & 0.3330 & 0.0799 \\ 0.0059 & 0.6907 & 0.0003 & 0.2534 \\ 0.0903 & 0.0218 & 0.6660 & 0.1599 \\ 0.0006 & 0.0691 & 0.0006 & 0.5068 \end{pmatrix}, \quad p_2^b(19) = 0.4922,$$

$$P_{\rho_1^{M_1,B_2}}(19) = \begin{pmatrix} 0.9023 & 0.2874 & 0.4992 & 0.1578 \\ 0.0068 & 0.6217 & 0.0008 & 0.3422 \\ 0.0902 & 0.0287 & 0.4992 & 0.1578 \\ 0.0007 & 0.0622 & 0.0008 & 0.3422 \end{pmatrix}, \quad p_1^b(19) = 0.3979.$$

If we examine each single plot from Figures 4.3 up to 4.6 then we we conclude that the points of the differences in subsequent norms tend to lie on a straight line. Notice that

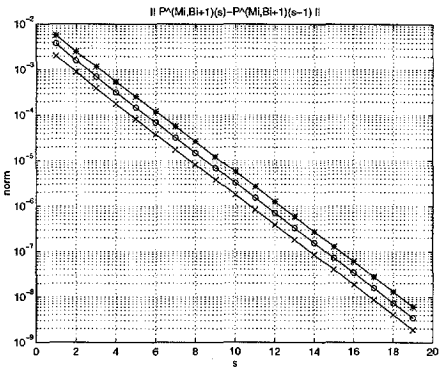

Figure 4.5: $\| P_{\rho_{i+1}^{M_i,B_{i+1}}}(s) - P_{\rho_{i+1}^{M_i,B_{i+1}}}(s-1) \|$ for i=1,2,3 we used "*","o","x" respectively.

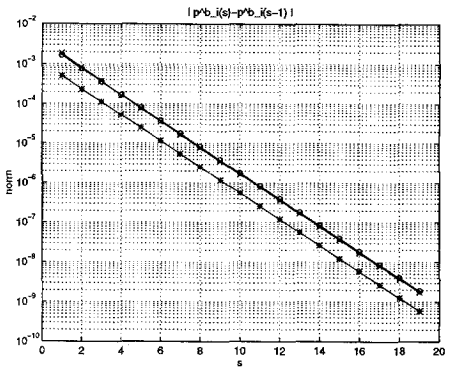

Figure 4.6: $\mid p_i^b(s) - p_i^b(s-1) \mid$ for i=1,2,3 we used "*","o","x" respectively.

if these points after some iteration steps lie on one monotonic descending line on the logarithmic scale then convergence of the iteration method is assured. Therefore, the plots show that from the first 19 iteration steps we may conclude that convergence is attained for all matrices and scalars involved in the iteration method.

After the numerical computations of the backward and forward aggregation steps we are also able by means of simple manipulations of the different stationary distribution vectors to make plots of the stationary distribution of the buffer contents for each separate buffer. By means of these plots the usage and performance of each buffer can be examined. The computed distribution of the contents of the buffers B_2, B_3 and B_4 can be denoted by $\tilde{\pi}_{\rho_2^{\text{Cont}}}$, $\tilde{\pi}_{\rho_3^{\text{Cont}}}$ and $\tilde{\pi}_{\rho_4^{\text{Cont}}}$. They are shown in the plots in Figures 4.7, 4.9 and 4.11 respectively.

Of course, the solution to the restrictions of our decomposed production line does *not* give the exact stationary distribution of the original production line. In order to evaluate the accuracy of the previous solution we have also done some simulations of the original production line. Therefore we have made a program for simulation of our production line in C programming language. We use again the same program as we have used for simulations earlier for the production lines with product failures only, we only changed the parameters such that they suit for our example here. The most important tool for the simulations is the random generator used. For our simulations in C we used again the *drand48* command which we have used also in examples from Chapter 2 and Chapter 3 in Section 2.8 and Section 3.11 respectively. The simulations are performed here in a similar way as we performed the simulations in the corresponding sections in Chapter 2 and Chapter 3. That means that again the pseudo-random number generator is initialized

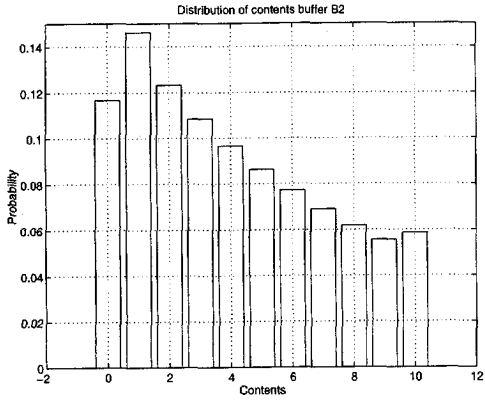

Figure 4.7: The *computed* stationary distribution of the contents of buffer B_2: $\tilde{\pi}_{\rho_2^{\text{Cont.}}}$.

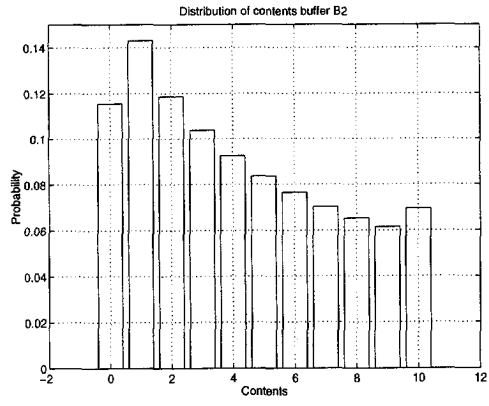

Figure 4.8: The *simulated* stationary distribution of the contents of buffer B_2.

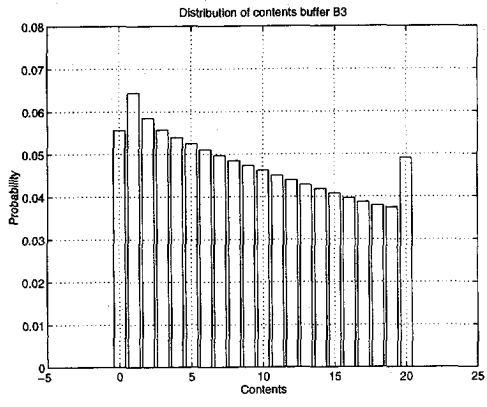

Figure 4.9: The *computed* stationary distribution of the contents of buffer B_3: $\tilde{\pi}_{\rho_3^{\text{Cont.}}}$.

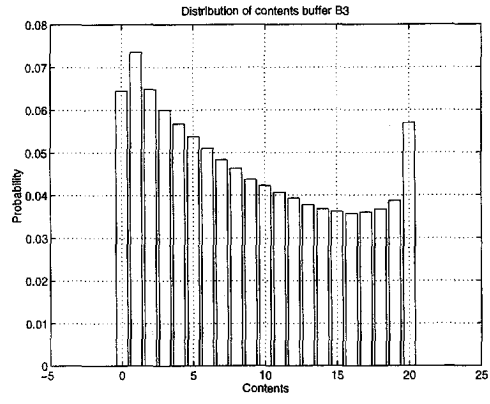

Figure 4.10: The *simulated* stationary distribution of the contents of buffer B_3.

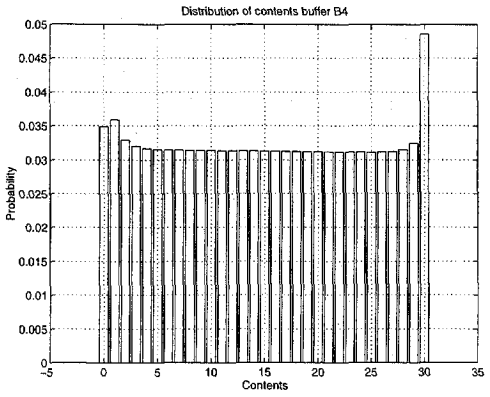

Figure 4.11: The *computed* stationary distribution of the contents of buffer B_4: $\tilde{\pi}_{\rho_4^{\text{Cont.}}}$.

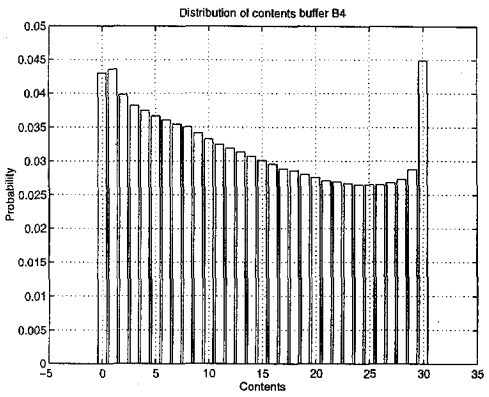

Figure 4.12: The *simulated* stationary distribution of the contents of buffer B_4.

Estimates of average production rate R by simulation										
	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
R	0.326933	0.326695	0.326830	0.326206	0.326701	0.327282	0.325595	0.326680	0.327027	0.327372
Estimates of the average buffer contents by simulation										
Buffer	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
B_2	4.1966	4.2264	4.2552	4.1548	4.2637	4.1873	4.2378	4.1224	4.1965	4.1603
B_3	8.7108	9.0105	8.9637	8.6200	9.0101	9.0029	9.0877	8.7750	8.8884	8.9243
B_4	14.0084	14.2376	13.9422	13.3544	14.1551	14.4926	14.0685	13.9570	13.6391	13.9521

Table 4.1: The results of 10^6 time steps simulations for the 4-machine-production line

by means of a time dependent argument so as to generate an independent sequence each time it is invoked.

The results with respect to the stationary distribution of the buffer contents of buffers B_2, B_3 and B_4 are shown in Figures 4.8, 4.10 and 4.12. These figures are put next to the figures of the computed stationary distribution of the buffer contents for comparison. Other results concerning average production rate and average buffer contents of the 10 independent "runs" of simulation are shown in Table 4.1. The compound results are shown in Table 4.2.

□

Results for the average production rate R						
	mean value	std. deviation	decomp. model		abs. err.	rel. err.
R	0.3267321	0.0005189	B_2	0.3296359	0.0029038	0.0088874
			B_3	0.3297002	0.0029681	0.0090842
			B_4	0.3297481	0.0030160	0.0092308

Results for the average buffer contents					
Buffer	mean value	std. deviation	decomp. model	abs. err.	rel. err.
B_2	4.2001	0.0460	4.0633	-0.1368	-0.0326
B_3	8.8993	0.1508	9.2000	0.3007	0.0338
B_4	13.9807	0.3125	15.1104	1.1297	0.0808

Table 4.2: The compound results of the simulations for the 4-machine-production line

4.7 Conclusions

In this chapter we described a decomposition method for a discrete synchronous serial production line with product part failures and machine failures. This means that the machines in the line have geometrically distributed service times and also that machines are prone to breakdown and repair. By means of the decomposition method presented here we are able to estimate the average production rate R of such production lines. The method constists of alternate forward and backward aggregation steps and has the same structure as the methods described in Chapter 2 and Chapter 3 for production lines with product failures only or production lines with machine failures only. In fact the method that we described in this chapter can be seen as an integration of the method that we described in Chapter 2 with the method that we described in Chapter 3. Because it was already impossible for the more simple method in Chapter 3, we have again not been able to prove convergence of the algorithm in this chapter and uniqueness of the solution. All examples however, of which one is treated here, show that the algorithm does converge and that the method is quite accurate.

Accuracy can be examined by comparison with the results from simulations. For the accuracy we distinguish between the estimates of the average production rate R and the estimates of the average contents of the separate buffers. The estimates of the average production rate R are within a range of 2 percent error normally, which we do consider as good. The estimates of the average contents of the buffers however, as we have already seen in Chapter 2 Section 2.8, can exceed 10 percent relative error for the worst cases.

Of course, since the decomposition method in this chapter is a direct combination of

the decomposition methods from the previous chapters, the method here cannot be more accurate than the methods there. Therefore our method here is at least as inaccurate as the method that we described in Chapter 2. However, it appears that the accuracy of the method in this chapter is *similar* to the accuracy of the method in Chapter 2, which seems reasonable because the method described in Chapter 3 was much more accurate.

Computations of the iteration method were done in the MATLAB software environment. In this environment the actual numerical computations such as 16×16 matrix inversion is quick and efficient. Inefficient however is the memory management of the MATLAB software. That is why MATLAB spends a lot of the execution time in data management tasks. The example treated in this section, just as the example treated in the previous chapter in Section 3.11, costs some minutes of computation time on a HP 720 workstation. We repeat that a tremendous speed up of the iteration method therefore could be obtained by translation of the method in a lower level programming language such as C. However, this was outside the scope of this research.

It is due to the decomposition method that there is a *linear* relation between the complexity of the algorithm and an increment in capacity of a buffer in the line. This property makes it possible to compute approximated versions of the stationary distribution even if the capacities of the buffers are large.

Chapter 5

An Application with Merging in a Car Lamp Factory

In this chapter we will describe how the method of alternate construction and reduction of Markov chains can be used in a car lamp production line that has a configuration with *merging*. That means that there is a machine that merges different product parts from two upstream buffers into one product part and puts the result in a single downstream buffer.

5.1 The Car Lamp Production Line

A schematic picture of the car lamp production line is shown in Figure 5.1. In this car lamp production line we have two machines that construct separately but parallel in time the inner metal parts and the outer glass bulbs of a car lamp. The inner metal parts, simply called *inner parts*, are produced by the so-called *mounting machine*. The outer bulb parts, simply called *bulbs*, are produced by the so-called *stemming machine*. The stemming machine and the mounting machine put the bulbs and inner parts in separate buffers which lead to the so-called *pump&pinch machine* that takes both, the inner and outer parts. In this pump-and-pinch machine the bulbs are pumped vacuum, refilled with a special gas, heated and then pinched with the metal parts inside. This machine pinches bulbs in order to close each bulb and hold the inner metal part to end up as a low voltage lamp that is able to give light, simply called a *burner*. The burners are again put in a buffer that leads to the so-called *testing machine* which tests if the burners have sufficient quality and puts a parabolic shaped reflecting cone around the burner for the finishing touch to a ready-for-use car lamp.

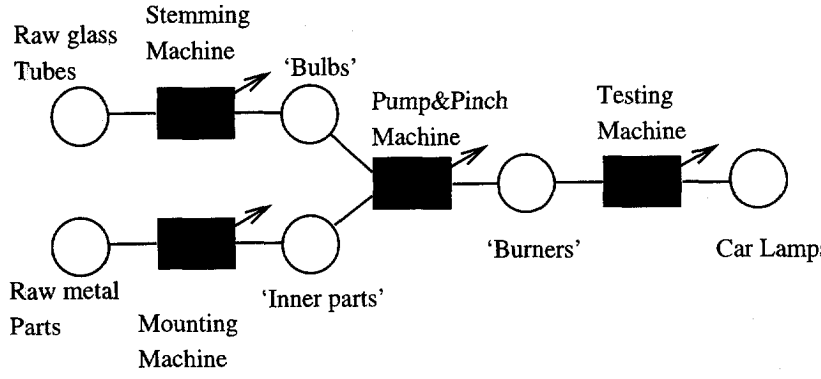

Figure 5.1: The car lamp production line schematically.

5.2 Description of the Car Lamp Production Line Model

The description of the model here is slightly different from the description of the models in previous chapters. The main differences between the model in the car lamp production line here and the previous models is the different handling of product part failures with the introduction of scrapping, and the new interpretation of the completion probability as the machine speed. The rest of the differences is due to the new configuration with merging which causes more administration during the computations.

5.2.1 A Formal Description of the Configuration

In order to avoid repetition of the long names of the machines and buffers we will rename them and give each machine and buffer a number or a combination of two numbers in case of machines or buffers that work in parallel. The new names of the machines and buffers and their number are listed in Table 5.1 together with the real names of the corresponding machines and buffers. As a result we obtain the translated version of the same production line as shown in Figure 5.2.

Formally, for the configuration of this car lamp production line we have a set of 4 machines denoted by $\mathcal{M}$, a set of 6 buffers denoted by $\mathcal{B}$ and a set of nine arcs denoted by $\mathcal{A}$ defined as follows:

$$\begin{aligned}
\mathcal{M} &\stackrel{\text{def}}{=} \{M_{1,1},\ M_{1,2},\ M_2,\ M_3\},\\
\mathcal{B} &\stackrel{\text{def}}{=} \{B_{1,1},\ B_{2,1},\ B_{1,2},\ B_{2,2},\ B_3,\ B_4\},\\
\mathcal{A} &\stackrel{\text{def}}{=} \{(B_{1,1}, M_{1,1}),\ (B_{1,2}, M_{1,2}),\ (B_{2,1}, M_2),\ (B_{2,2}, M_2),\ (B_3, M_3)\ldots\\
&\qquad\qquad (M_{1,1}, B_{2,1}),\ (M_{1,2}, B_{2,2}),\ (M_2, B_3),\ (M_3, B_4)\}.
\end{aligned}$$

Short name	Real name of machine or buffer
$M_{1,1}$	Stemming Machine
$M_{1,2}$	Mounting Machine
M_2	Pump&Pinch Machine
M_3	Testing Machine
$B_{1,1}$	Raw glass Tubes Buffer
$B_{2,1}$	Bulbs Buffer
$B_{1,2}$	Raw metal Parts Buffer
$B_{2,2}$	Inner parts Buffer
B_3	Burners Buffer
B_4	Car Lamps Buffer

Table 5.1: The short names and the corresponding real names in the car lamp production line.

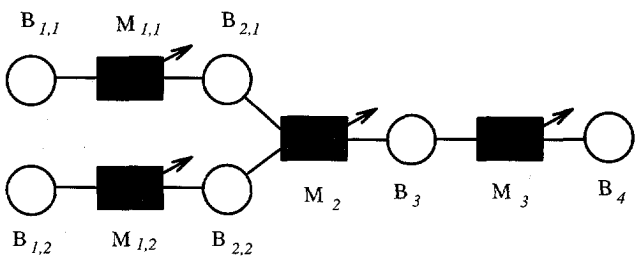

Figure 5.2: The car lamp production line with short names.

Buffers $B_{1,1}$ and $B_{1,2}$ are considered to be infinite capacity input buffers and buffer B_4 is considered to be an infinite capacity output buffer.

5.2.2 Description of the Machine Model

A difference with machine models in previous chapters is that in the previous models we assume that if no failures would occur then every machine produces exactly one product part per time slot. This would mean, apart from the failures, that in principle all machines produce parts with exactly the same speed. The different completion probabilities of separate machines however took care of the fact that the actual production per machine per time slot changed to different rates lower than one.

From completion probability to machine speed

To clarify how we will deal with the relation between production rate, speed and completion probability of a single machine we will introduce two different machines.

- Machine number *one* that attempts to manipulate product parts with a speed of exactly one product part per second. Suppose also that this machine one has a completion probability of $\frac{1}{2}$, which means that manipulation every time cycle fails with a probability of $\frac{1}{2}$. A product part that failed manipulation will be manipulated again in exactly the same manner in the next time cycle. The overall average production rate of this single machine one would be $\frac{1}{2}$ product parts per second.

- Machine number *two* that is not prone to failures, but produces product parts with a speed of exactly one part every two seconds. The actual average production rate of this single machine two would also be $\frac{1}{2}$ product parts per second.

Of course we realize that there are fundamental differences between the machines one and two that we just described, but on the level of average production rates they behave similarly. It is in this step that the model here is potentially liable to criticism. However, in order to be able to apply our Markov chain methods we need its convenient "memoryless" properties such that the same transitions occur with the same probabilities each time slot, which implies here that we can do three things

1. We study and construct more complex kinds of Markov chain models in which concurrent events with *variable durations* can take place. This means that we step from the idea that time is slotted in equal time slots and that we step from the idea that each time slot is similar with respect to the events and transitions that may occur.

2. We increase the number of states in our Markov chain models such that a recent part of the history is "included" and modeled as extra states in a new Markov chain model. This will keep our models based on time slots that are similar with respect to

size and with respect to the events and transitions that may occur. The procedure allows us also to make a clear distinction between machine number one and machine number two. The price that we pay for that is again an increase of the number of states in the state set.

3. We model machine number one by means of machine number two because this will still make sense on the level of average rates and continue as if this were correct.

In this chapter we have chosen for the *last* solution in 3. In the next chapter, Chapter 6, we will try to develop more complex Markov chain models as we just described in 1.

The choice that we just made for the solution in 3 implies that the notion *completion probability* changes to the notion *machine speed.* In other words, from now on in this chapter a completion probability $\frac{1}{2}$ does not imply that product manipulation fails with probability $\frac{1}{2}$, it only implies that it takes 2 time slots for the machine to produce one product part.

The time slot size

The relation between completion probabilities and machine speeds gives us the freedom to "choose" the size of the time slots for our overall Markov chain model. We can choose any size as long as the length of one time slot is smaller than the cycle time of the fastest machine $M_i \in \mathcal{M}$, such that all completion probabilities p_i will be smaller than 1. For the car lamp production line we notice that the fastest machine which is mounting machine $M_{1,2}$ that has a cycle time over 1 second. This is why we decided to choose for a time slot with a length of 1 second for all the following Markov chain models.

The product part failures and scrapping

Unfortunately, the fact that the completion probability p_i of a machine $M_i \in \mathcal{M}$ does not relate to product part failures does not mean that product part failures do not occur in this machine. Product part failures do occur for machine M_i, but the way that the machine handles with these failures is different from the models that we used in the previous chapters. In the previous chapters if a product part manipulation failure occurred in one time slot we tried to finish manipulation in the next time slot with the same failure probability. If a product part manipulation failure occurs in a machine M_i of the car lamp production line here then the product part will immediately be thrown away in a *garbage bin* and will *not* return in the line. This way of dealing with failures of product part manipulations is called *scrapping.* Since the completion probability p_i now deals with machine speed only we introduce a new parameter c_i that deals with the product part failures with scrapping for machine M_i. This new parameter will be called the *scrap probability* of M_i and denotes the probability that machine M_i fails manipulation and throws away a certain product part. We will assume that the same scrap probability c_i

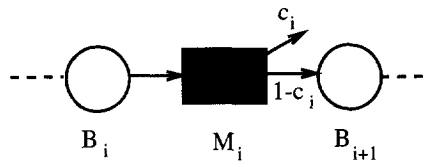

Figure 5.3: A machine M_i with scrap probability c_i.

of a machine M_i applies to every product part that passes M_i for manipulation, totally independent of time and totally independent of the state of other machines and buffers in the production line. The scrapping of product parts by a machine will be depicted symbolically by means of an extra diagonal downstream arc from the the machine that does not point at a downstream buffer. A picture of a machine with scrapping is shown in Figure 5.3.

The breakdown-repair model

The breakdown-repair model that we use for the separate machines $M_i \in \mathcal{M}$ in the car lamp production line is exactly the same as the breakdown-repair model that we used in the serial production lines with machine and product part failures combined in the previous chapter. The only difference with the model in the previous chapter is the interpretation of the completion probability as the machine speed. We assume that a machine M_i can only be in two different states: it is either *up* or *down*.

When up, the machine is in progress and manipulates product parts with a completion rate, i.e. a speed, given by the completion probability related to the machine. When down, the machine is not in progress and does not manipulate product parts. The completion probability, i.e. speed, in down state evidently equals 0.

The transitions between the states up and down are organized as follows:
When up, the machine has a certain probability to go down during the current time slot. This probability is called the breakdown rate α_i. If the machine does not break down in the current time slot it will be up when the next time slot starts. When the machine does break down during the current time slot, it will be down at the beginning of the next time slot.
When down, the machine has a certain probability to go up during the current time slot. This probability is called the repair rate β_i. If the machine will not be repaired during the current time slot it will be down when the next time slot starts. When the machine is repaired during the current time slot, it will be up at the beginning of the next time slot. The simple Markov chain corresponding to the breakdown repair model is shown in Figure 5.4.

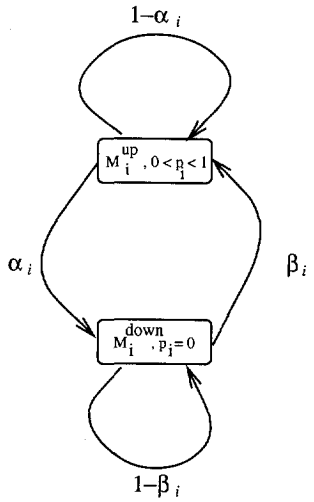

Figure 5.4: The breakdown-repair model of a machine M_i.

Simple calculus shows that the average consecutive up time (in time slots) is as follows:

$$\begin{aligned} \mathbb{E}[\text{ UP time } M_i \text{ }] &= \alpha_i \sum_{j=0}^{\infty} j\ (1-\alpha_i)^j \\ &= \frac{1-\alpha_i}{\alpha_i}\ . \end{aligned} \tag{5.1}$$

In exactly the same way the average consecutive down time can be derived:

$$\begin{aligned} \mathbb{E}[\text{ DOWN time } M_i \text{ }] &= \beta_i \sum_{j=0}^{\infty} j\ (1-\beta_i)^j \\ &= \frac{1-\beta_i}{\beta_i}\ . \end{aligned} \tag{5.2}$$

Suppose one wants to cunstruct such a breakdown-repair Markov chain model given real data about breakdown and repair times in practice. Then we can use Equations (5.1) and (5.2) to determine the breakdown and repair rates α_i and β_i. The procedure is very simple: First compute (in time slots) the average consecutive up time as well as the average consecutive down time by means of the data. The result will be denoted by x_i and y_i respectively. Then compute, as follows from Expressions (5.1) and (5.2), the breakdown and repair rates α_i and β_i:

$$\alpha_i = \frac{1}{1+x_i}\ ,$$

$$\beta_i \;=\; \frac{1}{1+y_i}\,. \tag{5.3}$$

The breakdown-repair process for one single machine M_i can be described by the following simple Markov chain:

$$\pi(t+1) = \begin{pmatrix} 1-\alpha_i & \beta_i \\ \alpha_i & 1-\beta_i \end{pmatrix} \pi(t)\,. \tag{5.4}$$

The stationary distribution π of this Markov chain is:

$$\pi = \frac{1}{\alpha_i + \beta_i} \begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix}. \tag{5.5}$$

5.3 Decomposition of the Car Lamp Production Line: Three Different Buffer Models with Sub-Models

In the car lamp production line we deal with three finite capacity buffers, buffer $B_{2,1}$, buffer $B_{2,2}$ and buffer B_3. In order to make computation less complicated we decompose the total car lamp production line model into three overlapping sub-models. Each sub-model corresponds to one of the buffers with finite capacity and its environment which consists of the closest up- and downstream machines and buffers. The way of reasoning for each of these buffer models is similar to the way of reasoning in Chapter 3, especially Section 3.3, except that here we have for each buffer model an extra buffer in the environment. This extra buffer can be in two extra states which implies that the number of states of each buffer environment here is twice the number of states in the environment of the buffer models in Chapter 3. The description of the three different buffer models starts naturally with the description of the states inside each of the models. For the description of the states we will assume that the finite capacities of buffers $B_{2,1}$, $B_{2,2}$ and B_3 are $N_{2,1} \in I\!N^+$, $N_{2,2} \in I\!N^+$ and $N_3 \in I\!N^+$ respectively. Just as in Chapter 3 Section 3.3 we will, for the sake of convenience of notation, first define the following state sets of separate machines and buffers:

$$\begin{array}{rcl@{\qquad}rcl}
\mathcal{S}^{\text{Contents } B_{2,1}} & \stackrel{\text{def}}{=} & \{0,1,..,N_{2,1}\}\,, & \mathcal{S}^{\text{Starving } B_{1,2}} & \stackrel{\text{def}}{=} & \{B_{1,2}^{\neg\text{st}}, B_{1,2}^{\text{st}}\}\,, \\
\mathcal{S}^{\text{Contents } B_{2,2}} & \stackrel{\text{def}}{=} & \{0,1,..,N_{2,2}\}\,, & \mathcal{S}^{\text{Starving } B_{2,2}} & \stackrel{\text{def}}{=} & \{B_{2,2}^{\neg\text{st}}, B_{2,2}^{\text{st}}\}\,, \\
\mathcal{S}^{\text{Contents } B_3} & \stackrel{\text{def}}{=} & \{0,1,..,N_3\}\,, & \mathcal{S}^{\text{State } M_{1,1}} & \stackrel{\text{def}}{=} & \{M_{1,1}^{\text{up}}, M_{1,1}^{\text{down}}\}\,, \\
\mathcal{S}^{\text{Blocking } B_3} & \stackrel{\text{def}}{=} & \{B_3^{\neg\text{bl}}, B_3^{\text{bl}}\}\,, & \mathcal{S}^{\text{State } M_{1,2}} & \stackrel{\text{def}}{=} & \{M_{1,2}^{\text{up}}, M_{1,2}^{\text{down}}\}\,, \\
\mathcal{S}^{\text{Blocking } B_4} & \stackrel{\text{def}}{=} & \{B_4^{\neg\text{bl}}, B_4^{\text{bl}}\}\,, & \mathcal{S}^{\text{State } M_2} & \stackrel{\text{def}}{=} & \{M_2^{\text{up}}, M_2^{\text{down}}\}\,, \\
\mathcal{S}^{\text{Starving } B_{1,1}} & \stackrel{\text{def}}{=} & \{B_{1,1}^{\neg\text{st}}, B_{1,1}^{\text{st}}\}\,, & \mathcal{S}^{\text{State } M_3} & \stackrel{\text{def}}{=} & \{M_3^{\text{up}}, M_3^{\text{down}}\}\,, \\
\mathcal{S}^{\text{Starving } B_{2,1}} & \stackrel{\text{def}}{=} & \{B_{2,1}^{\neg\text{st}}, B_{2,1}^{\text{st}}\}\,. & & & \\
\end{array} \tag{5.6}$$

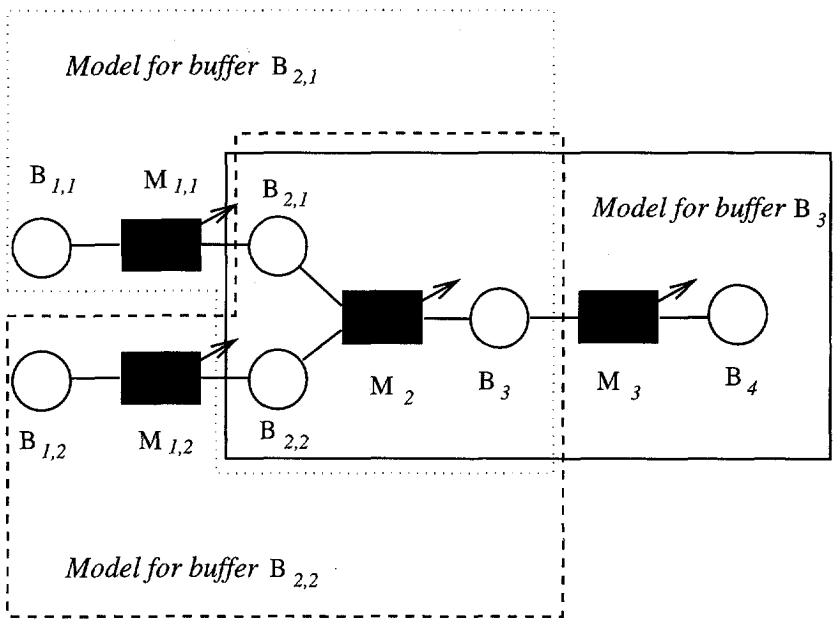

Figure 5.5: The three sub-models of the car lamp production line.

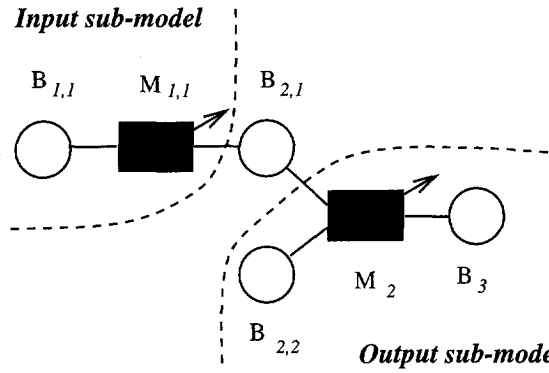

Figure 5.6: Buffer $B_{2,1}$ and its environment: the input sub-model and the output sub-model.

The decomposition in sub-models is shown in Figure 5.5.

In the following we will describe the three different models each in more detail.

5.3.1 The States of the Model for Buffer $B_{2,1}$ and the Input and Output Sub-Model

Buffer $B_{2,1}$ and its environment form the basis of the model for buffer $B_{2,1}$. A picture is shown in Figure 5.6. For this model for buffer $B_{2,1}$ we define state sets $\mathcal{S}^{B_{2,1}}$ and $\mathcal{R}^{B_{2,1}}$ as follows:

$$\mathcal{S}^{B_{2,1}} \stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{1,1}} \times \mathcal{S}^{\text{State } M_{1,1}} \times \mathcal{S}^{\text{Contents } B_{2,1}} \times \mathcal{S}^{\text{Starving } B_{2,2}} \times \mathcal{S}^{\text{Blocking } B_3} \times \mathcal{S}^{\text{State } M_2},$$

$$\mathcal{R}^{B_{2,1}} \stackrel{\text{def}}{=} \{0, 1, 2, ..., N_{2,1}\} \times \{1, 2, ..., 32\}.$$

As defined, the state set $\mathcal{R}^{B_{2,1}}$ consists of ordered pairs of numbers of which the first number corresponds to the contents of the buffer and the second number corresponds to one of the 32 states of the environment. In Table C.1 in Appendix C we show the one to one correspondence between the states in set $\mathcal{S}^{B_{2,1}}$ and the states in set $\mathcal{R}^{B_{2,1}}$, similar to the table in Table 3.1. Inside this model for buffer $B_{2,1}$ we can again distinguish two sub-models. As shown in Figure 5.6, we will call these sub-models the *input sub-model* and the *output sub-model* of the model for buffer $B_{2,1}$. Both models do also have a state set which we will describe in more detail in the next two subsections.

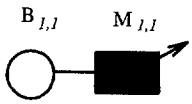

Figure 5.7: The input sub-model of the model for buffer $B_{2,1}$.

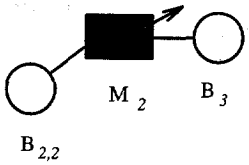

Figure 5.8: The output sub-model of the model for buffer $B_{2,1}$.

The states of the input sub-model of the model for buffer $B_{2,1}$

The input sub-model of the model is a model based on the combination of input buffer $B_{1,1}$ and machine $M_{1,1}$. A picture of this combination is shown in Figure 5.7.

$$\begin{aligned} \mathcal{S}^{B_{2,1},\text{in}} &\stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{1,1}} \times \mathcal{S}^{\text{State } M_{1,1}}, \\ \mathcal{R}^{B_{2,1},\text{in}} &\stackrel{\text{def}}{=} \{1,2,3,4\}. \end{aligned}$$

The one to one relation between states in $\mathcal{S}^{B_{2,1},\text{in}}$ and states in $\mathcal{R}^{B_{2,1},\text{in}}$ can be found in Table C.2.

The states of the output sub-model of the model for buffer $B_{2,1}$

The output sub-model of the model is a model based on the combination of buffer $B_{2,2}$, buffer B_3 and machine M_2. A picture of this combination is shown in Figure 5.8.

$$\begin{aligned} \mathcal{S}^{B_{2,1},\text{out}} &\stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{2,2}} \times \mathcal{S}^{\text{Blocking } B_3} \times \mathcal{S}^{\text{State } M_2}, \\ \mathcal{R}^{B_{2,1},\text{out}} &\stackrel{\text{def}}{=} \{1,2,3,4,5,6,7,8\}. \end{aligned}$$

The one to one relation between states in $\mathcal{S}^{B_{2,1},\text{out}}$ and states in $\mathcal{R}^{B_{2,1},\text{out}}$ can be found in Table C.3.

5.3.2 The States of the Model for Buffer $B_{2,2}$ and the Input and Output Sub-Model

Buffer $B_{2,2}$ and its environment form the basis of the model for buffer $B_{2,2}$. A picture is shown in Figure 5.9. For this model for buffer $B_{2,2}$ we define state sets $\mathcal{S}^{B_{2,2}}$ and $\mathcal{R}^{B_{2,2}}$ as

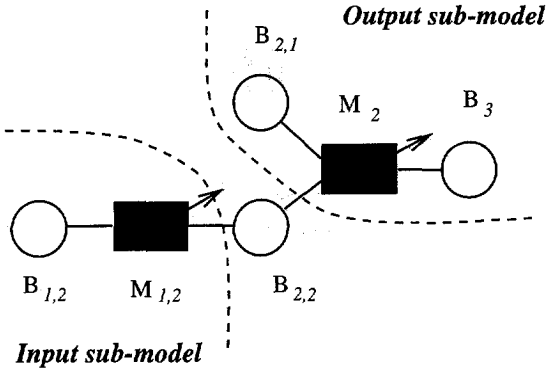

Figure 5.9: Buffer $B_{2,2}$ and its environment: the input sub-model and the output sub-model.

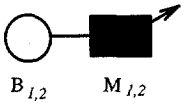

Figure 5.10: The input sub-model of the model for buffer $B_{2,2}$.

follows:

$$\mathcal{S}^{B_{2,2}} \stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{1,2}} \times \mathcal{S}^{\text{State } M_{1,2}} \times \mathcal{S}^{\text{Contents } B_{2,2}} \times \mathcal{S}^{\text{Starving } B_{2,1}} \times \mathcal{S}^{\text{Blocking } B_3} \times \mathcal{S}^{\text{State } M_2},$$

$$\mathcal{R}^{B_{2,2}} \stackrel{\text{def}}{=} \{0, 1, 2, ..., N_{2,2}\} \times \{1, 2, ..., 32\}.$$

As defined, the state set $\mathcal{R}^{B_{2,2}}$ consists of ordered pairs of numbers of which the first number corresponds to the contents of the buffer and the second number corresponds to one of the 32 states of the environment. In Table C.4 we show the one to one correspondence between the states in set $\mathcal{S}^{B_{2,2}}$ and the states in set $\mathcal{R}^{B_{2,2}}$, similar to the table in Table 3.1. Inside this model for buffer $B_{2,2}$ we can again distinguish two sub-models. As shown in Figure 5.9, we will call these sub-models the *input sub-model* and the *output sub-model* of the model for buffer $B_{2,2}$. Both models do also have a state set which we will describe in more detail in the next two subsections.

The states of the input sub-model of the model for buffer $B_{2,2}$

The input sub-model of the model is a model based on the combination of input buffer $B_{1,2}$ and machine $M_{1,2}$. A picture of this combination is shown in Figure 5.10.

$$\mathcal{S}^{B_{2,2},\text{in}} \stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{1,2}} \times \mathcal{S}^{\text{State } M_{1,2}},$$

$$\mathcal{R}^{B_{2,2},\text{in}} \stackrel{\text{def}}{=} \{1, 2, 3, 4\}.$$

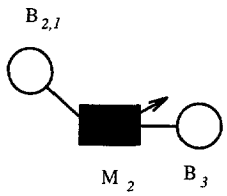

Figure 5.11: The output sub-model of the model for buffer $B_{2,2}$.

The one to one relation between states in $\mathcal{S}^{B_{2,2},\text{in}}$ and states in $\mathcal{R}^{B_{2,2},\text{in}}$ can be found in Table C.5.

The states of the output sub-model of the model for buffer $B_{2,2}$

The output sub-model of the model is a model based on the combination of buffer $B_{2,1}$, buffer B_3 and machine M_2. A picture of this combination is shown in Figure 5.11.

$$\begin{aligned}\mathcal{S}^{B_{2,2},\text{out}} &\stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{2,1}} \times \mathcal{S}^{\text{Blocking } B_3} \times \mathcal{S}^{\text{State } M_2},\\ \mathcal{R}^{B_{2,2},\text{out}} &\stackrel{\text{def}}{=} \{1,2,3,4,5,6,7,8\}.\end{aligned}$$

The one to one relation between states in $\mathcal{S}^{B_{2,2},\text{out}}$ and states in $\mathcal{R}^{B_{2,2},\text{out}}$ can be found in Table C.6.

5.3.3 The States of the Model for Buffer B_3 and the Input and Output Sub-Model

Buffer B_3 and its environment forms the basis of the model for buffer B_3. A picture is shown in Figure 5.12. For this model for buffer B_3 we define state sets $\mathcal{S}^{B_3}$ and $\mathcal{R}^{B_3}$ as follows:

$$\begin{aligned}\mathcal{S}^{B_3} &\stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{1,2}} \times \mathcal{S}^{\text{Starving } B_{2,2}} \times \mathcal{S}^{\text{State } M_2} \times \mathcal{S}^{\text{Contents } B_3} \times \mathcal{S}^{\text{Blocking } B_4} \times \mathcal{S}^{\text{State } M_3},\\ \mathcal{R}^{B_3} &\stackrel{\text{def}}{=} \{0,1,2,...,N_3\} \times \{1,2,...,32\}.\end{aligned}$$

As defined, the state set $\mathcal{R}^{B_3}$ consists of ordered pairs of numbers of which the first number corresponds to the contents of the buffer and the second number corresponds to one of the 32 states of the environment. In Table C.7 we show the one to one correspondence between the states in set $\mathcal{S}^{B_3}$ and the states in set $\mathcal{R}^{B_3}$, similar to the table in Table 3.1. Inside this model for buffer B_3 we can again distinguish two sub-models. As shown in Figure 5.12, we will call these sub-models the *input sub-model* and the *output sub-model* of the model for buffer B_3. Both models do also have a state set which we will describe in more detail in the next two subsections.

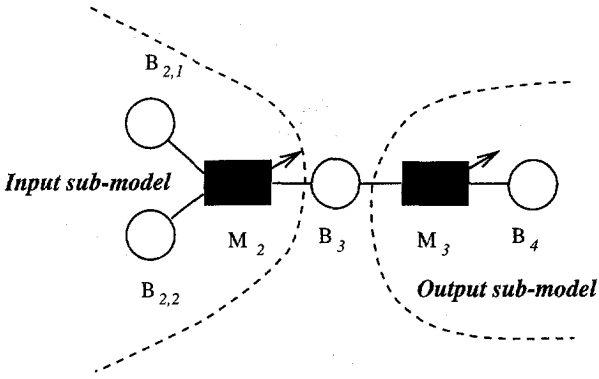

Figure 5.12: Buffer B_3 and its environment: the input sub-model and the output sub-model.

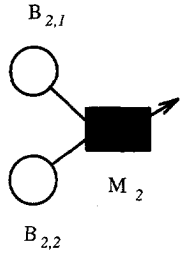

Figure 5.13: The input sub-model of the model for buffer B_3.

The states of the input sub-model of the model for buffer B_3

The input sub-model of the model is a model based on the combination of input buffer $B_{1,2}$, buffer $B_{2,2}$ and machine $M_{1,2}$. A picture of this combination is shown in Figure 5.13.

$$\begin{aligned} \mathcal{S}^{B_3,\text{in}} &\stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{2,1}} \times \mathcal{S}^{\text{Starving } B_{2,2}} \times \mathcal{S}^{\text{State } M_2}, \\ \mathcal{R}^{B_3,\text{in}} &\stackrel{\text{def}}{=} \{1,2,3,4,5,6,7,8\}. \end{aligned}$$

The one to one relation between states in $\mathcal{S}^{B_3,\text{in}}$ and states in $\mathcal{R}^{B_3,\text{in}}$ can be found in Table C.8.

The states of the output sub-model of the model for buffer B_3

The output sub-model of the model is a model based on the combination of buffer B_4 and machine M_3. A picture of this combination is shown in Figure 5.14.

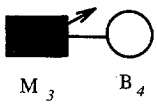

Figure 5.14: The output sub-model of the model for buffer B_3.

$$\mathcal{S}^{B_3,\text{out}} \stackrel{\text{def}}{=} \mathcal{S}^{\text{Blocking } B_4} \times \mathcal{S}^{\text{State } M_3},$$
$$\mathcal{R}^{B_3,\text{out}} \stackrel{\text{def}}{=} \{1,2,3,4\}.$$

The one to one relation between states in $\mathcal{S}^{B_3,\text{out}}$ and states in $\mathcal{R}^{B_3,\text{out}}$ can be found in Table C.9.

5.3.4 States of the Model for the Overlap Between the Buffer Models

In Figure 5.5 we can see that the three buffer models overlap. The overlap of the buffer models consists of the buffers $B_{2,1}$, $B_{2,2}$, B_3 and machine M_2. For this combination of machines and buffers, which we will call the *overlap* from now on, we will make a separate model as well. We distinguish two state sets of the overlap as follows:

$$\mathcal{S}^{\text{overlap}} \stackrel{\text{def}}{=} \mathcal{S}^{\text{Starving } B_{2,1}} \times \mathcal{S}^{\text{Starving } B_{2,2}} \times \mathcal{S}^{\text{Blocking } B_3} \times \mathcal{S}^{\text{State } M_2},$$
$$\mathcal{R}^{\text{overlap}} \stackrel{\text{def}}{=} \{1,2,3,...,15,16\}.$$

The one to one relation between the states in $\mathcal{S}^{\text{overlap}}$ and states in $\mathcal{R}^{\text{overlap}}$ can be found in Table C.10.

5.4 The Relations Between the States of the Various Models

In the previous section we described the state sets of three buffer models each with an input and an output sub-model. Next to these buffer models with sub-models we also described the state set of the overlap model. Obviously, the models that we described have strong relations with each another. The relations between the models are such that the states of a smaller model or sub-model represent groups of states from bigger models. This means that the Markov chain for such a bigger model can be reduced by means of a partitioning, as is described extensively in Chapter 1 Section 1.3, in order to obtain the Markov chain for the smaller model. Next in this section we will describe in detail the partitionings that we need for the bigger models in order to reduce them to one of the smaller models.

First we will start with the most obvious partitionings, the partitionings of states in the big buffer models that correspond with the states of their input and output sub-models.

5.4.1 The Clusters in $\mathcal{R}^{B_{2,1}}$ for Reduction to $\mathcal{R}^{B_{2,1},\text{in}}$

For the relation between the states of the model for $B_{2,1}$ and the states of the input sub-model for $B_{2,1}$ we define the following partitioning $\rho_{2,1}^{B_{2,1},\text{in}} \in \mathcal{P}(\mathcal{R}^{B_{2,1}})$.

$$\rho_{2,1}^{B_{2,1},\text{in}} \stackrel{\text{def}}{=} (\mathcal{R}_{2,1}^{(B_{1,1}^{\neg\text{st}},M_{1,1}^{\text{up}})}, \mathcal{R}_{2,1}^{(B_{1,1}^{\text{st}},M_{1,1}^{\text{up}})}, \mathcal{R}_{2,1}^{(B_{1,1}^{\neg\text{st}},M_{1,1}^{\text{down}})}, \mathcal{R}_{2,1}^{(B_{1,1}^{\text{st}},M_{1,1}^{\text{down}})}),$$

where

$$\begin{aligned}
\mathcal{R}_{2,1}^{(B_{1,1}^{\neg\text{st}},M_{1,1}^{\text{up}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{1,2,...,7,8\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{1,1}^{\text{st}},M_{1,1}^{\text{up}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{9,10,...,15,16\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{1,1}^{\neg\text{st}},M_{1,1}^{\text{down}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{17,18,...,23,24\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{1,1}^{\text{st}},M_{1,1}^{\text{down}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{25,26,...,31,32\}\ \}\ .
\end{aligned} \tag{5.7}$$

5.4.2 The Clusters in $\mathcal{R}^{B_{2,1}}$ for Reduction to $\mathcal{R}^{B_{2,1},\text{out}}$

For the relation between states of the model for $B_{2,1}$ and the states of the output sub-model for $B_{2,1}$ we define the following partitioning $\rho_{2,1}^{B_{2,1},\text{out}} \in \mathcal{P}(\mathcal{R}^{B_{2,1}})$.

$$\begin{aligned}
\rho_{2,1}^{B_{2,1},\text{out}} \stackrel{\text{def}}{=} (&\mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\neg\text{bl}},M_2^{\text{up}})}, \mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\neg\text{bl}},M_2^{\text{down}})}, \mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\text{bl}},M_2^{\text{up}})}, ... \\
&...\mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\text{bl}},M_2^{\text{down}})}, \mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\neg\text{bl}},M_2^{\text{up}})}, \mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\neg\text{bl}},M_2^{\text{down}})}, ... \\
&...\mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\text{bl}},M_2^{\text{up}})}, \mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\text{bl}},M_2^{\text{down}})}),
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\neg\text{bl}},M_2^{\text{up}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{1,9,17,25\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\neg\text{bl}},M_2^{\text{down}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{2,10,18,26\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\text{bl}},M_2^{\text{up}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{3,11,19,27\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\neg\text{st}},B_3^{\text{bl}},M_2^{\text{down}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{4,12,20,28\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\neg\text{bl}},M_2^{\text{up}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{5,13,21,29\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\neg\text{bl}},M_2^{\text{down}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{6,14,22,30\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\text{bl}},M_2^{\text{up}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{7,15,23,31\}\ \}\ , \\
\mathcal{R}_{2,1}^{(B_{2,2}^{\text{st}},B_3^{\text{bl}},M_2^{\text{down}})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid b \in \{8,16,24,32\}\ \}.
\end{aligned} \tag{5.8}$$

5.4.3 The Clusters in $\mathcal{R}^{B_{2,2}}$ for Reduction to $\mathcal{R}^{B_{2,2},\text{in}}$

For the relation between the states of the model for $B_{2,2}$ and the states of the input sub-model for $B_{2,2}$ we define the following partitioning $\rho_{2,2}^{B_{2,2},\text{in}} \in \mathcal{P}(\mathcal{R}^{B_{2,2}})$.

$$\rho_{2,2}^{B_{2,2},\text{in}} \stackrel{\text{def}}{=} (\mathcal{R}_{2,2}^{(B_{1,2}^{\neg\text{st}},M_{1,2}^{\text{up}})}, \mathcal{R}_{2,2}^{(B_{1,2}^{\text{st}},M_{1,2}^{\text{up}})}, \mathcal{R}_{2,2}^{(B_{1,2}^{\neg\text{st}},M_{1,2}^{\text{down}})}, \mathcal{R}_{2,2}^{(B_{1,2}^{\text{st}},M_{1,2}^{\text{down}})}),$$

where

$$\begin{aligned}
\mathcal{R}_{2,2}^{(B_{1,2}^{\neg st}, M_{1,2}^{up})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{1,2,...,7,8\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{1,2}^{st}, M_{1,2}^{up})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{9,10,...,15,16\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{1,2}^{\neg st}, M_{1,2}^{down})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{17,18,...,23,24\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{1,2}^{st}, M_{1,2}^{down})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{25,26,...,31,32\}\ \}\ .
\end{aligned} \tag{5.9}$$

5.4.4 The Clusters in $\mathcal{R}^{B_{2,2}}$ for Reduction to $\mathcal{R}^{B_{2,2},\text{out}}$

For the relation between the states of the model for $B_{2,2}$ and the states of the output sub-model for $B_{2,2}$ we define the following partitioning $\rho_{2,2}^{B_{2,2},\text{out}} \in \mathcal{P}(\mathcal{R}^{B_{2,2}})$.

$$\begin{aligned}
\rho_{2,2}^{B_{2,2},\text{out}} \stackrel{\text{def}}{=} (&\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{\neg bl}, M_2^{up})}, \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{\neg bl}, M_2^{down})}, \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{bl}, M_2^{up})}, ... \\
&...\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{bl}, M_2^{down})}, \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{\neg bl}, M_2^{up})}, \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{\neg bl}, M_2^{down})}, ... \\
&...\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{bl}, M_2^{up})}, \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{bl}, M_2^{down})}),
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{\neg bl}, M_2^{up})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{1,9,17,25\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{\neg bl}, M_2^{down})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{2,10,18,26\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{bl}, M_2^{up})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{3,11,19,27\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_3^{bl}, M_2^{down})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{4,12,20,28\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{\neg bl}, M_2^{up})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{5,13,21,29\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{\neg bl}, M_2^{down})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{6,14,22,30\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{bl}, M_2^{up})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{7,15,23,31\}\ \}\ , \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_3^{bl}, M_2^{down})} &= \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid b \in \{8,16,24,32\}\ \}.
\end{aligned} \tag{5.10}$$

5.4.5 The Clusters in $\mathcal{R}^{B_3}$ for Reduction to $\mathcal{R}^{B_3,\text{in}}$

For the relation between the states of the model for B_3 and the states of the input sub-model for B_3 we define the following partitioning $\rho_{2,2}^{B_3,\text{in}} \in \mathcal{P}(\mathcal{R}^{B_3})$.

$$\begin{aligned}
\rho_3^{B_3,\text{in}} \stackrel{\text{def}}{=} (&\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, M_2^{down})}, \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, M_2^{up})}, ... \\
&...\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, M_2^{down})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, M_2^{down})}, ... \\
&...\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, M_2^{down})}),
\end{aligned}$$

where

$$
\begin{aligned}
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, M_2^{up})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{1,2,3,4\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, M_2^{down})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{5,6,7,8\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, M_2^{up})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{9,10,11,12\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, M_2^{down})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{13,14,15,16\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, M_2^{up})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{17,18,19,20\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, M_2^{down})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{21,22,23,24\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, M_2^{up})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{25,26,27,28\} \,\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, M_2^{down})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{29,30,31,32\} \,\}.
\end{aligned}
\tag{5.11}
$$

5.4.6 The Clusters in $\mathcal{R}^{B_3}$ for Reduction to $\mathcal{R}^{B_3,\text{out}}$

For the relation between the states of the model for B_3 and the states of the output sub-model for B_3 we define the following partitioning $\rho_{2,2}^{B_3,\text{out}} \in \mathcal{P}(\mathcal{R}^{B_3})$.

$$
\rho_3^{B_3,\text{out}} \stackrel{\text{def}}{=} (\mathcal{R}_3^{(B_4^{\neg bl}, M_3^{up})}, \mathcal{R}_3^{(B_4^{\neg bl}, M_3^{down})}, \mathcal{R}_3^{(B_4^{bl}, M_3^{up})}, \mathcal{R}_3^{(B_4^{bl}, M_3^{down})}),
$$

where

$$
\begin{aligned}
\mathcal{R}_3^{(B_4^{\neg bl}, M_3^{up})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{1,5,9,13,17,21,25,29\} \,\}, \\
\mathcal{R}_3^{(B_4^{\neg bl}, M_3^{down})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{2,6,10,14,18,22,26,30\} \,\}, \\
\mathcal{R}_3^{(B_4^{bl}, M_3^{up})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{3,7,11,15,19,23,27,31\} \,\}, \\
\mathcal{R}_3^{(B_4^{bl}, M_3^{down})} &= \{\, (a,b) \in \mathcal{R}^{B_3} \mid b \in \{4,8,12,16,20,24,28,32\} \,\}.
\end{aligned}
\tag{5.12}
$$

The next partitionings are more difficult and more important. These partitionings reduce the states of the three buffer models in clusters that correspond to the states of the model for the overlap.

5.4.7 The Clusters in $\mathcal{R}^{B_{2,1}}$ for Reduction to $\mathcal{R}^{\text{overlap}}$

For the relation between the states of the model for $B_{2,1}$ and the states of the overlap model we define the following partitioning $\rho_{2,1}^{\text{overlap}} \in \mathcal{P}(\mathcal{R}^{B_{2,1}})$.

$$
\begin{aligned}
\rho_{2,1}^{\text{overlap}} \stackrel{\text{def}}{=} (&\mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})}, \ldots
\end{aligned}
$$

$$
\begin{array}{l}
\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{up}})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{down}})}, \ldots \\
\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{up}})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{down}})}, \ldots \\
\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{up}})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{down}})}, \ldots \\
\ldots \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{up}})}, \quad \mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{down}})}),
\end{array}
$$

where

$$
\begin{aligned}
\mathcal{R}_{2,1}^{(B_{2,1}^{\neg \mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{1,9,17,25\}) \bigvee (a = 0 \bigwedge b = 1) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\neg \mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{2,10,18,26\}) \bigvee (a = 0 \bigwedge b = 2) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\neg \mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{3,11,19,27\}) \bigvee (a = 0 \bigwedge b = 3) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\neg \mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{4,12,20,28\}) \bigvee (a = 0 \bigwedge b = 4) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\neg \mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{5,13,21,29\}) \bigvee (a = 0 \bigwedge b = 5) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\neg \mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{6,14,22,30\}) \bigvee (a = 0 \bigwedge b = 6) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{7,15,23,31\}) \bigvee (a = 0 \bigwedge b = 7) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{8,16,24,32\}) \bigvee (a = 0 \bigwedge b = 8) \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{9,17,25\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{10,18,26\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{11,19,27\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\neg \mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{12,20,28\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{13,21,29\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\neg \mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{14,22,30\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{up}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{15,23,31\} \}, \\
\mathcal{R}_{2,1}^{(B_{2,1}^{\mathrm{st}}, B_{2,2}^{\mathrm{st}}, B_3^{\mathrm{bl}}, M_2^{\mathrm{down}})} &= \{ (a,b) \in \mathcal{R}^{B_{2,1}} \mid a = 0 \bigwedge b \in \{16,24,32\} \}.
\end{aligned}
\tag{5.13}
$$

5.4.8 The Clusters in $\mathcal{R}^{B_{2,2}}$ for Reduction to $\mathcal{R}^{\text{overlap}}$

For the relation between the states of the model for $B_{2,2}$ and the states of the overlap model we define the following partitioning $\rho_{2,2}^{\text{overlap}} \in \mathcal{P}(\mathcal{R}^{B_{2,2}})$.

$$
\begin{aligned}
\rho_{2,2}^{\text{overlap}} \stackrel{\text{def}}{=} (&\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})}, \ldots \\
&\ldots \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})}, \ \mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{1,9,17,25\}) \bigvee (a = 0 \bigwedge b = 1) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{2,10,18,26\}) \bigvee (a = 0 \bigwedge b = 2) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{3,11,19,27\}) \bigvee (a = 0 \bigwedge b = 3) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{4,12,20,28\}) \bigvee (a = 0 \bigwedge b = 4) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{9,17,25\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{10,18,26\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{11,19,27\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{12,20,28\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{5,13,21,29\}) \bigvee (a = 0 \bigwedge b = 5) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{6,14,22,30\}) \bigvee (a = 0 \bigwedge b = 6) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{7,15,23,31\}) \bigvee (a = 0 \bigwedge b = 7) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{8,16,24,32\}) \bigvee (a = 0 \bigwedge b = 8) \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{13,21,29\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{14,22,30\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{15,23,31\} \}, \\
\mathcal{R}_{2,2}^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})} &= \{ (a,b) \in \mathcal{R}^{B_{2,2}} \mid a = 0 \bigwedge b \in \{16,24,32\} \}.
\end{aligned}
\tag{5.14}
$$

5.4.9 The Clusters in $\mathcal{R}^{B_3}$ for Reduction to $\mathcal{R}^{\text{overlap}}$

For the relation between the states of the model for B_3 and the states of the overlap model we define the following partitioning $\rho_3^{\text{overlap}} \in \mathcal{P}(\mathcal{R}^{B_3})$.

$$
\begin{aligned}
\rho_3^{\text{overlap}} \stackrel{\text{def}}{=} (&\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})}, \dots \\
&\dots \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})}, \mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{1,2,3,4\}) \bigvee (a = N_3 \bigwedge b = 1)\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{5,6,7,8\}) \bigvee (a = N_3 \bigwedge b = 5)\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{2,3,4\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{6,7,8\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{9,10,11,12\}) \bigvee (a = N_3 \bigwedge b = 9)\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{13,14,15,16\}) \bigvee (a = N_3 \bigwedge b = 13)\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{10,11,12\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{\neg st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{14,15,16\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{17,18,19,20\}) \bigvee (a = N_3 \bigwedge b = 17)\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{\neg bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{21,22,23,24\}) \bigvee (a = N_3 \bigwedge b = 21)\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{18,19,20\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{\neg st}, B_3^{bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{22,23,24\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{25,26,27,28\}) \bigvee (a = N_3 \bigwedge b = 25)\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{\neg bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ (a < N_3 \bigwedge b \in \{29,30,31,32\}) \bigvee (a = N_3 \bigwedge b = 29)\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{up})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{26,27,28\}\}, \\
\mathcal{R}_3^{(B_{2,1}^{st}, B_{2,2}^{st}, B_3^{bl}, M_2^{down})} &= \{(a,b) \in \mathcal{R}^{B_3} |\ a = N_3 \bigwedge b \in \{30,31,32\}\}.
\end{aligned}
\tag{5.15}
$$

Next we have the more straightforward relations between clusters of states of the overlap and the states of the two output sub-models for buffers $B_{2,1}$ and $B_{2,2}$ and the relation between the clusters of states of the overlap and the states of the input sub-model for buffer B_3.

5.4.10 The Clusters in $\mathcal{R}^{\text{overlap}}$ for Reduction to $\mathcal{R}^{B_{2,1},\text{out}}$

For the relation between the states of the overlap model and the states of the output sub-model for $B_{2,1}$ we define the following partitioning $\rho^{B_{2,1},\text{out}}_{\text{overlap}} \in \mathcal{P}(\mathcal{R}^{\text{overlap}})$.

$$
\begin{aligned}
\rho^{B_{2,1},\text{out}}_{\text{overlap}} \stackrel{\text{def}}{=} (&\mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \ldots \\
&\ldots\mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}, \ldots \\
&\ldots\mathcal{R}^{(B^{\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{1,9\} \,\}, \\
\mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{2,10\} \,\}, \\
\mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{3,11\} \,\}, \\
\mathcal{R}^{(B^{\neg\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{4,12\} \,\}, \\
\mathcal{R}^{(B^{\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{5,13\} \,\}, \\
\mathcal{R}^{(B^{\text{st}}_{2,2},B^{\neg\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{6,14\} \,\}, \\
\mathcal{R}^{(B^{\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{7,15\} \,\}, \\
\mathcal{R}^{(B^{\text{st}}_{2,2},B^{\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{8,16\} \,\}.
\end{aligned}
\tag{5.16}
$$

5.4.11 The Clusters in $\mathcal{R}^{\text{overlap}}$ for Reduction to $\mathcal{R}^{B_{2,2},\text{out}}$

For the relation between the states of the overlap model and the states of the output sub-model for $B_{2,2}$ we define the following partitioning $\rho^{B_{2,2},\text{out}}_{\text{overlap}} \in \mathcal{P}(\mathcal{R}^{\text{overlap}})$.

$$
\begin{aligned}
\rho^{B_{2,2},\text{out}}_{\text{overlap}} \stackrel{\text{def}}{=} (&\mathcal{R}^{(B^{\neg\text{st}}_{2,1},B^{\neg\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\neg\text{st}}_{2,1},B^{\neg\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\neg\text{st}}_{2,1},B^{\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \ldots \\
&\ldots\mathcal{R}^{(B^{\neg\text{st}}_{2,1},B^{\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\text{st}}_{2,1},B^{\neg\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\text{st}}_{2,1},B^{\neg\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}, \ldots \\
&\ldots\mathcal{R}^{(B^{\text{st}}_{2,1},B^{\text{bl}}_{3},M^{\text{up}}_{2})}_{\text{overlap}}, \mathcal{R}^{(B^{\text{st}}_{2,1},B^{\text{bl}}_{3},M^{\text{down}}_{2})}_{\text{overlap}}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_3^{\neg\text{bl}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{1,5\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_3^{\neg\text{bl}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{2,6\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_3^{\text{bl}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{3,7\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_3^{\text{bl}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{4,8\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_3^{\neg\text{bl}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{9,13\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_3^{\neg\text{bl}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{10,14\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_3^{\text{bl}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{11,15\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_3^{\text{bl}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{12,16\} \,\}.
\end{aligned}
\tag{5.17}
$$

5.4.12 The Clusters in $\mathcal{R}^{\text{overlap}}$ for Reduction to $\mathcal{R}^{B_3,\text{in}}$

For the relation between the states of the overlap model and the states of the input submodel for B_3 we define the following partitioning $\rho_{\text{overlap}}^{B_3,\text{in}} \in \mathcal{P}(\mathcal{R}^{\text{overlap}})$.

$$
\begin{aligned}
\rho_{\text{overlap}}^{B_3,\text{in}} \stackrel{\text{def}}{=} (&\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{up}})}, \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{down}})}, \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\text{st}},M_2^{\text{up}})}, \ldots \\
&\ldots \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\text{st}},M_2^{\text{down}})}, \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{up}})}, \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{down}})}, \ldots \\
&\ldots \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\text{st}},M_2^{\text{up}})}, \mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\text{st}},M_2^{\text{down}})}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{1,3\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{2,4\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\text{st}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{5,7\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\neg\text{st}},B_{2,2}^{\text{st}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{6,8\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{9,11\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\neg\text{st}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{10,12\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\text{st}},M_2^{\text{up}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{13,15\} \,\}, \\
\mathcal{R}_{\text{overlap}}^{(B_{2,1}^{\text{st}},B_{2,2}^{\text{st}},M_2^{\text{down}})} &= \{\, a \in \mathcal{R}^{\text{overlap}} \mid a \in \{14,16\} \,\}.
\end{aligned}
\tag{5.18}
$$

5.5 Construction of the Markov Chains for the Three Buffer Models

In this section we will describe how we construct the Markov chains for the three buffer models. The construction of each of the Markov chain models for a buffer is based on the

assumption that both the process corresponding to the input sub-model and the process corresponding to the output sub-model are independent Markov processes. Next we will describe the construction of each of the three buffer models: the model for buffer $B_{2,1}$, the model for buffer $B_{2,2}$ and the model for buffer B_3.

5.5.1 Construction of the Markov Chain of the Model for Buffer $B_{2,1}$

First we define five diagonal matrices of size 32, denoted by $\Lambda_A^{2,1}$, $\Lambda_B^{2,1}$, $\Lambda_C^{2,1}$ $\Lambda_D^{2,1}$ and $\Lambda_E^{2,1}$, each as a function of two scalars $p_{1,1}, p_2 \in [0,1]$ with $q_{1,1} = 1 - p_{1,1}$ and $q_2 = 1 - p_2$ as follows:

$$
\begin{aligned}
\Lambda_A^{2,1}(p_{1,1},p_2) &\stackrel{\text{def}}{=} \begin{array}{l}\text{diag}(q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1}, \quad 1,1,1,1,1,1,1,1,\ldots \\ \quad \ldots 1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ , \quad 1,1,1,1,1,1,1,1\ ),\end{array} \\
\Lambda_B^{2,1}(p_{1,1},p_2) &\stackrel{\text{def}}{=} \begin{array}{l}\text{diag}(p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1}, \quad 0,0,0,0,0,0,0,0,\ldots \\ \quad \ldots 0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ , \quad 0,0,0,0,0,0,0,0\ ),\end{array} \\
\Lambda_C^{2,1}(p_{1,1},p_2) &\stackrel{\text{def}}{=} \begin{array}{l}\text{diag}(p_2q_{1,1},0,0,0,0,0,0,0, \quad p_2,0,0,0,0,0,0,0,\ldots \\ \quad \ldots p_2,0,0,0,0,0,0,0, \quad p_2,0,0,0,0,0,0,0\ ),\end{array} \\
\Lambda_D^{2,1}(p_{1,1},p_2) &\stackrel{\text{def}}{=} \\
&\begin{array}{l}\text{diag}(p_{1,1}p_2+q_{1,1}q_2,q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1},q_{1,1}, \quad q_2,1,1,1,1,1,1,1,\ldots \\ \quad \ldots q_2 \quad ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ , \quad q_2,1,1,1,1,1,1,1\ ),\end{array} \\
\Lambda_E^{2,1}(p_{1,1},p_2) &\stackrel{\text{def}}{=} \begin{array}{l}\text{diag}(p_{1,1}q_2,p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1},p_{1,1}, \quad 0,0,0,0,0,0,0,0,\ldots \\ \quad \ldots 0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ , \quad 0,0,0,0,0,0,0,0\ ).\end{array}
\end{aligned}
\tag{5.19}
$$

Suppose that the Markov transition matrix of the input sub-model for buffer $B_{2,1}$ is equal to the 4×4 stochastic matrix X. Suppose that the Markov transition matrix of the output sub-model for buffer $B_{2,1}$ is equal to the 8×8 stochastic matrix Y. Suppose also that machine $M_{1,1}$, in case it is up and buffer $B_{1,1}$ is not starved, puts x product parts per time slot in buffer $B_{2,1}$. Suppose that machine M_2, in case it is up and buffer $B_{2,2}$ is not starved and buffer B_3 is not blocked, takes y product parts per time slot from buffer $B_{2,1}$. Under these circumstances, and under the assumption that the input and output sub-processes are totally independent, the Markov transition matrix $P^{B_{2,1}}$ of the model for buffer $B_{2,1}$ as a function of x, y, X and Y can be formulated as follows:

$$
\begin{aligned}
P^{B_{2,1}}(x,y,X,Y) &\stackrel{\text{def}}{=} \\
&\left[I_{N_{2,1}} \otimes X \otimes Y\right] \begin{pmatrix} \Lambda_A^{2,1}(x,y) & \Lambda_C^{2,1}(x,y) & & & \\ \Lambda_B^{2,1}(x,y) & \Lambda_D^{2,1}(x,y) & \Lambda_C^{2,1}(x,y) & & \\ & \ddots & \ddots & \ddots & \\ & & \Lambda_E^{2,1}(x,y) & \Lambda_D^{2,1}(x,y)+\Lambda_E^{2,1}(x,y) \end{pmatrix}.
\end{aligned}
\tag{5.20}
$$

5.5.2 Construction of the Markov Chain of the Model for Buffer $B_{2,2}$

First we define five diagonal matrices of size 32, denoted by $\Lambda_A^{2,2}$, $\Lambda_B^{2,2}$, $\Lambda_C^{2,2}$ $\Lambda_D^{2,2}$ and $\Lambda_E^{2,2}$, each as a function of two scalars $p_{1,2}, p_2 \in [0,1]$ with $q_{1,2} = 1 - p_{1,2}$ and $q_2 = 1 - p_2$ as follows:

$$
\begin{aligned}
\Lambda_A^{2,2}(p_{1,2},p_2) &\stackrel{\text{def}}{=} \operatorname{diag}(q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2}, \quad 1,1,1,1,1,1,1,1,\ldots \\
&\qquad \ldots 1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ , \quad 1,1,1,1,1,1,1,1), \\
\Lambda_B^{2,2}(p_{1,2},p_2) &\stackrel{\text{def}}{=} \operatorname{diag}(p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2}, \quad 0,0,0,0,0,0,0,0,\ldots \\
&\qquad \ldots 0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ , \quad 0,0,0,0,0,0,0,0), \\
\Lambda_C^{2,2}(p_{1,2},p_2) &\stackrel{\text{def}}{=} \operatorname{diag}(p_2q_{1,2},0,0,0,0,0,0,0, \quad p_2,0,0,0,0,0,0,0,\ldots \\
&\qquad \ldots p_2,0,0,0,0,0,0,0, \quad p_2,0,0,0,0,0,0,0), \\
\Lambda_D^{2,2}(p_{1,2},p_2) &\stackrel{\text{def}}{=} \\
&\operatorname{diag}(p_{1,2}p_2+q_{1,2}q_2,q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2},q_{1,2}, \quad q_2,1,1,1,1,1,1,1,\ldots \\
&\qquad \ldots q_2 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ ,1 \ , \quad q_2,1,1,1,1,1,1,1), \\
\Lambda_E^{2,2}(p_{1,2},p_2) &\stackrel{\text{def}}{=} \operatorname{diag}(p_{1,2}q_2,p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2},p_{1,2}, \quad 0,0,0,0,0,0,0,0,\ldots \\
&\qquad \ldots 0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ ,0 \ , \quad 0,0,0,0,0,0,0,0).
\end{aligned}
\tag{5.21}
$$

Suppose that the Markov transition matrix of the input sub-model for buffer $B_{2,2}$ is equal to the 4×4 stochastic matrix X. Suppose that the Markov transition matrix of the output sub-model for buffer $B_{2,2}$ is equal to the 8×8 stochastic matrix Y. Suppose also that machine $M_{1,2}$, in case it is up and buffer $B_{1,2}$ is not starved, puts x product parts per time slot in buffer $B_{2,2}$. Suppose that machine M_2, in case it is up and buffer $B_{2,2}$ is not starved and buffer B_3 is not blocked, takes y product parts per time slot from buffer $B_{2,2}$. Under these circumstances, and under the assumption that the input and output sub-processes are totally independent, the Markov transition matrix $P^{B_{2,2}}$ of the model for buffer $B_{2,2}$ as a function of x, y, X and Y can be formulated as follows:

$$
\begin{aligned}
P^{B_{2,2}}(x,y,X,Y) &\stackrel{\text{def}}{=} \\
&\left[I_{N_{2,2}} \otimes X \otimes Y\right] \begin{pmatrix} \Lambda_A^{2,2}(x,y) & \Lambda_C^{2,2}(x,y) & & & \\ \Lambda_B^{2,2}(x,y) & \Lambda_D^{2,2}(x,y) & \Lambda_C^{2,2}(x,y) & & \\ & \ddots & \ddots & \ddots & \\ & & \Lambda_E^{2,2}(x,y) & \Lambda_D^{2,2}(x,y)+\Lambda_E^{2,2}(x,y) \end{pmatrix}.
\end{aligned}
\tag{5.22}
$$

5.5.3 Construction of the Markov Chain of the Model for Buffer B_3

First we define five diagonal matrices of size 32, denoted by Λ^3_A, Λ^3_B, Λ^3_C Λ^3_D and Λ^3_E, each as a function of two scalars $p_2, p_3 \in [0,1]$ with $q_3 = 1 - p_3$ and $q_3 = 1 - p_3$ as follows:

$$\begin{aligned}
\Lambda^3_A(p_2,p_3) &\stackrel{\text{def}}{=} \text{diag}(q_2,q_2,q_2,q_2,\ 1,1,1,1,\ 1,1,1,1,\ 1,1,1,1,... \\
&\qquad ...1,1,1,1,\ 1,1,1,1,\ 1,1,1,1,\ 1,1,1,1), \\
\Lambda^3_B(p_2,p_3) &\stackrel{\text{def}}{=} \text{diag}(p_2,p_2,p_2,p_2,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0,... \\
&\qquad ...0,0,0,0,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0), \\
\Lambda^3_C(p_2,p_3) &\stackrel{\text{def}}{=} \text{diag}(p_3q_2,0,0,0,\ p_3,0,0,0,\ p_3,0,0,0,\ p_3,0,0,0,... \\
&\qquad ...p_3,0,0,0,\ p_3,0,0,0,\ p_3,0,0,0,\ p_3,0,0,0), \\
\Lambda^3_D(p_2,p_3) &\stackrel{\text{def}}{=} \text{diag}(p_2p_3+q_2q_3,q_2,q_2,q_2,\ q_3,1,1,1,\ q_3,1,1,1,\ q_3,1,1,1,... \\
&\qquad ...q_3\ ,1,1,1,\ q_3,1,1,1,\ q_3,1,1,1,\ q_3,1,1,1), \\
\Lambda^3_E(p_2,p_3) &\stackrel{\text{def}}{=} \text{diag}(p_2q_3,p_2,p_2,p_2,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0,... \\
&\qquad ...0,0,0,0,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0).
\end{aligned} \tag{5.23}$$

Suppose that the Markov transition matrix of the input sub-model for buffer B_3 is equal to the 8×8 stochastic matrix X. Suppose that the Markov transition matrix of the output sub-model for buffer B_3 is equal to the 4×4 stochastic matrix Y. Suppose also that machine M_2, in case it is up and both buffers $B_{2,1}$ and $B_{2,2}$ are not starved, puts x product parts per time slot in buffer B_3. Suppose that machine M_3, in case it is up and buffer B_4 is not blocked, takes y product parts per time slot from buffer B_3. Under these circumstances, and under the assumption that the input and output sub-processes are totally independent, the Markov transition matrix P^{B_3} of the model for buffer B_3 as a function of x, y, X and Y can be formulated as follows:

$$\begin{aligned}
P^{B_3}(x,y,X,Y) &\stackrel{\text{def}}{=} \\
&[I_{N_3} \otimes X \otimes Y] \begin{pmatrix}
\Lambda^3_A(x,y) & \Lambda^3_C(x,y) & & & \\
\Lambda^3_B(x,y) & \Lambda^3_D(x,y) & \Lambda^3_C(x,y) & & \\
& \ddots & \ddots & \ddots & \\
& & \Lambda^3_E(x,y) & \Lambda^3_D(x,y)+\Lambda^3_E(x,y)
\end{pmatrix}.
\end{aligned} \tag{5.24}$$

5.6 Linkage of the Three Buffer Models

In this section we describe the relations between the three buffer models for buffers $B_{2,1}$, $B_{2,2}$ and B_3. The relations between the three buffer models can be divided into two kinds: long term relations and short term relations. The long term relations describe the long term phenomena such as blocking and starvation of buffers and breakdown and repair of machines. The short term relations describe the short term phenomena such as machine speed, scrapping, short term blocking and short term starvation. Next we will describe the long and short term relations between the three buffer models separately.

5.6.1 The Long Term Relations Between the Three Buffer Models

The most natural restriction to make for the linkage of the three buffer models is the restriction that reduction of the Markov chains of each of the three buffer models to a Markov chain for the overlap model results in models for the overlap that are equal. In terms of equations, if we denote the transition matrix of the Markov chain of the overlap by P^{overlap}, we can formulate that as follows:

$$P^{\text{overlap}} = \tilde{P}^{B_{2,1}}_{\rho^{\text{overlap}}_{2,1}} = \tilde{P}^{B_{2,2}}_{\rho^{\text{overlap}}_{2,2}} = \tilde{P}^{B_3}_{\rho^{\text{overlap}}_{3}}. \tag{5.25}$$

This restriction should also be consistent with the construction of the Markov chains for buffers $B_{2,1}$, $B_{2,2}$ and B_3 by means of the input and output sub-models. If we denote the Markov transition matrices corresponding to the input and output sub-models for buffers $B_{2,1}$, $B_{2,2}$ and B_3 by $P^{B_{2,1},\text{in}}$, $P^{B_{2,1},\text{out}}$, $P^{B_{2,2},\text{in}}$, $P^{B_{2,2},\text{out}}$ and $P^{B_3,\text{in}}$, $P^{B_3,\text{out}}$ then we can formulate that as follows:

$$\begin{aligned} P^{B_{2,1}} &= P^{B_{2,1}}(x_{2,1}, y_{2,1}, P^{B_{2,1},\text{in}}, P^{B_{2,1},\text{out}}), \\ P^{B_{2,2}} &= P^{B_{2,2}}(x_{2,2}, y_{2,2}, P^{B_{2,2},\text{in}}, P^{B_{2,2},\text{out}}), \\ P^{B_3} &= P^{B_3}(x_3, y_3, P^{B_3,\text{in}}, P^{B_3,\text{out}}), \end{aligned} \tag{5.26}$$

where $x_{2,1}$, $x_{2,2}$, x_3 stand for the numbers of parts per time slot that enter, and $y_{2,1}$, $y_{2,2}$, y_3 stand for the numbers of parts per time slot that leave the buffers $B_{2,1}$, $B_{2,2}$, B_3 respectively. The last parameters adjust for instance for the different speeds and the scrapping of the separate machines. We will return to the subject of fine-tuning the model with these parameters in the next subsection on short term relations.

Next to these restrictions we have the restriction that the Markov chain model for the overlap can be reduced in three ways. Each way corresponds with one of the following Markov chains: the Markov chain of the output sub-model for buffer $B_{2,1}$, the Markov chain of the output sub-model for buffer $B_{2,2}$ or the Markov chain of the input sub-model for buffer B_3:

$$\begin{aligned} P^{B_{2,1},\text{out}} &= \tilde{P}^{\text{overlap}}_{\rho^{B_{2,1},\text{out}}_{\text{overlap}}}, \\ P^{B_{2,2},\text{out}} &= \tilde{P}^{\text{overlap}}_{\rho^{B_{2,2},\text{out}}_{\text{overlap}}}, \\ P^{B_3,\text{in}} &= \tilde{P}^{\text{overlap}}_{\rho^{B_3,\text{in}}_{\text{overlap}}}. \end{aligned} \tag{5.27}$$

5.6.2 The Short Term Relations Between the Three Buffer Models

In this subsection we discuss the short term relations between the three buffer models, which means that we focus on situations in which machines are up and buffers are not long term starved or blocked. For the description of the short term relations we need the stationary distributions of the three buffer models for buffers $B_{2,1}$, $B_{2,2}$ and B_3 which we will denote by $\pi^{B_{2,1}}$, $\pi^{B_{2,2}}$ and π^{B_3} respectively. Therefore, by definition, the following relations hold:

$$\begin{aligned} P^{B_{2,1}}\ \pi^{B_{2,1}} &= \pi^{B_{2,1}}, \\ P^{B_{2,2}}\ \pi^{B_{2,2}} &= \pi^{B_{2,2}}, \\ P^{B_3}\ \pi^{B_3} &= \pi^{B_3}. \end{aligned} \tag{5.28}$$

By means of these stationary distributions we can derive the short term starvation and short term blocking probabilities. Here we will focus on three of these probabilities: the short term starvation probability of buffer $B_{2,1}$, the short term starvation probability of buffer $B_{2,2}$ and the short term blocking probability of buffer B_3. These probabilities, as a function of the corresponding transition matrix, will be denoted by $\Psi_{\text{st}}^{B_{2,1}}(P^{B_{2,1}})$, $\Psi_{\text{st}}^{B_{2,2}}(P^{B_{2,2}})$ and $\Psi_{\text{bl}}^{B_3}(q_3, P^{B_3})$. For the definitions of these probabilities we will use the following subsets that restrict the system to short term situations in which machines are up and buffers are not long term starved or long term blocked:

$$\mathcal{R}_{2,1}^{(B_{2,1}^{\neg\text{st}}, B_{2,2}^{\neg\text{st}}, B_3^{\neg\text{bl}}, M_2^{\text{up}})} = \{\ (a,b) \in \mathcal{R}^{B_{2,1}} \mid (a > 0 \bigwedge b \in \{1,9,17,25\}) \bigvee (a = 0 \bigwedge b = 1)\}\ ,$$

$$\mathcal{R}_{2,2}^{(B_{2,1}^{\neg\text{st}}, B_{2,2}^{\neg\text{st}}, B_3^{\neg\text{bl}}, M_2^{\text{up}})} = \{\ (a,b) \in \mathcal{R}^{B_{2,2}} \mid (a > 0 \bigwedge b \in \{1,9,17,25\}) \bigvee (a = 0 \bigwedge b = 1)\}\ ,$$

$$\mathcal{R}_{3}^{(B_{2,1}^{\neg\text{st}}, B_{2,2}^{\neg\text{st}}, B_3^{\neg\text{bl}}, M_2^{\text{up}})} = \{\ (a,b) \in \mathcal{R}^{B_3} \mid (a < N_3 \bigwedge b \in \{1,2,3,4\}) \bigvee (a = N_3 \bigwedge b = 1)\}\ .$$

These subsets are equal to the subsets that we have used in Expressions (5.13), (5.14) and (5.15). The definitions are as follows:

$$\begin{aligned} \Psi_{\text{st}}^{B_{2,1}}(P^{B_{2,1}}) &\stackrel{\text{def}}{=} \frac{\pi_1^{B_{2,1}}}{\sum_{(a,b)\in\mathcal{R}_{2,1}^{(B_{2,1}^{\neg\text{st}}, B_{2,2}^{\neg\text{st}}, B_3^{\neg\text{bl}}, M_2^{\text{up}})}} \pi_{32a+b}^{B_{2,1}}}, \\ \Psi_{\text{st}}^{B_{2,2}}(P^{B_{2,2}}) &\stackrel{\text{def}}{=} \frac{\pi_1^{B_{2,2}}}{\sum_{(a,b)\in\mathcal{R}_{2,2}^{(B_{2,1}^{\neg\text{st}}, B_{2,2}^{\neg\text{st}}, B_3^{\neg\text{bl}}, M_2^{\text{up}})}} \pi_{32a+b}^{B_{2,2}}}, \\ \Psi_{\text{bl}}^{B_3}(q_3, P^{B_3}) &\stackrel{\text{def}}{=} \frac{q_3 \cdot \pi_{32N_3+1}^{B_3}}{\sum_{(a,b)\in\mathcal{R}_{3}^{(B_{2,1}^{\neg\text{st}}, B_{2,2}^{\neg\text{st}}, B_3^{\neg\text{bl}}, M_2^{\text{up}})}} \pi_{32a+b}^{B_3}}. \end{aligned} \tag{5.29}$$

By means of these short term starvation and blocking probabilities $\Psi_{\text{st}}^{B_{2,1}}$, $\Psi_{\text{st}}^{B_{2,2}}$ and $\Psi_{\text{bl}}^{B_3}$ we can more easily formulate the short term relations that we will use between the three buffer models:

$$\begin{aligned} x_{2,1} &= p_{1,1} \cdot (1 - c_{1,1}) \ , \\ x_{2,2} &= p_{1,2} \cdot (1 - c_{1,2}) \ , \\ y_3 &= p_3 \ , \\ y_{2,1} &= p_2 \cdot (1 - \Psi_{\text{st}}^{B_{2,2}}) \cdot (1 - \Psi_{\text{bl}}^{B_3}) \ , \\ y_{2,2} &= p_2 \cdot (1 - \Psi_{\text{st}}^{B_{2,1}}) \cdot (1 - \Psi_{\text{bl}}^{B_3}) \ , \\ x_3 &= p_2 \cdot (1 - c_2) \cdot (1 - \Psi_{\text{st}}^{B_{2,2}}) \cdot (1 - \Psi_{\text{st}}^{B_{2,1}}) \ . \end{aligned} \tag{5.30}$$

(Note that for convenience of notation here we have omitted the arguments of the functions $\Psi_{\text{st}}^{B_{2,1}}(P^{B_{2,1}})$, $\Psi_{\text{st}}^{B_{2,2}}(P^{B_{2,2}})$ and $\Psi_{\text{bl}}^{B_3}(q_3, P^{B_3})$.)
We translate the equations in Expression (5.30) as follows:

- The number of product parts per time slot that enter buffer $B_{2,1}$ ($x_{2,1}$) equals the number of manipulated product parts by machine $M_{1,1}$ per time slot ($p_{1,1}$) times the fraction of product parts that remains after testing and scrapping ($1 - c_{1,1}$).

- The number of product parts per time slot that enter buffer $B_{2,2}$ ($x_{2,2}$) equals the number of manipulated product parts by machine $M_{1,1}$ per time slot ($p_{1,2}$) times the fraction of product parts that remains after testing and scrapping ($1 - c_{1,2}$).

- The number of product parts per time slot that leave buffer B_3 (y_3) is simply determined by the speed (p_3) of machine M_3.

- A product part leaves buffer $B_{2,1}$ with the speed of machine M_2 (p_2). This speed ($y_{2,1}$) however is corrected and multiplied by the fraction of time that buffer $B_{2,2}$ is not short term starved ($1 - \Psi_{\text{st}}^{B_{2,2}}$) and multiplied by the fraction of time that buffer B_3 is not short term blocked ($1 - \Psi_{\text{bl}}^{B_3}$).

- A product part leaves buffer $B_{2,2}$ with the speed of machine M_2 (p_2). This speed ($y_{2,2}$) however is corrected and multiplied by the fraction of time that buffer $B_{2,1}$ is not short term starved ($1 - \Psi_{\text{st}}^{B_{2,1}}$) and multiplied by the fraction of time that buffer B_3 is not short term blocked ($1 - \Psi_{\text{bl}}^{B_3}$).

- The number of product parts per time slot that enter buffer B_3 (x_3) equals the number of manipulated product parts by machine M_2 per time slot (p_2) times the fraction of product parts that remains after testing and scrapping ($1 - c_2$). This number however should also be corrected and multiplied by the fraction of time that buffer $B_{2,1}$ is not short term starved ($1 - \Psi_{\text{st}}^{B_{2,1}}$) and multiplied by the fraction of time that buffer $B_{2,2}$ is not short term starved ($1 - \Psi_{\text{st}}^{B_{2,2}}$).

5.7 The Average Input and Output Rates in the Three Buffer Models

In order to determine the average production rate of the car lamp production line it is important to have insight in how the average input and output of each buffer can be computed by means of the corresponding buffer model.

Note for the determination of the average input and output rates that, for each of the three buffer models, there are states with a conflict. The model for buffer $B_{2,1}$ has conflict states $(0,1) \in \mathcal{R}^{B_{2,1}}$ and $(N_{2,1},1) \in \mathcal{R}^{B_{2,1}}$. The model for buffer $B_{2,2}$ has conflict states $(0,1) \in \mathcal{R}^{B_{2,2}}$ and $(N_{2,2},1) \in \mathcal{R}^{B_{2,2}}$. The model for buffer B_3 has conflict states $(0,1) \in \mathcal{R}^{B_3}$ and $(N_3,1) \in \mathcal{R}^{B_3}$. In each of these conflict states the order in which the up- and downstream machines operate determines the next state in the buffer model. In these conflict situations we will assume that the downstream machine always tries to take a product part before the upstream machine tries to put one.

5.7.1 The Average Input and Output Rate of Buffer $B_{2,1}$

Input in buffer $B_{2,1}$ can only occur in one of the following two cases:

1. Buffer $B_{2,1}$ contains less than $N_{2,1}$ product parts, buffer $B_{1,1}$ is not starved, machine $M_{1,1}$ is up and produces a (non-scrapped) product part.

2. Buffer $B_{2,1}$ contains exactly $N_{2,1}$ product parts, buffer $B_{1,1}$ is not starved, machine $M_{1,1}$ is up and produces a (non-scrapped) product part *and* buffer $B_{2,2}$ is not starved, buffer B_3 is not blocked, machine M_2 is up and takes a product part.

The probability that the first case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 1}) = x_{2,1} \cdot \Big(\sum_{\substack{(a,b) \in \mathcal{R}^{B_{2,1}} \\ a < N_{2,1} \\ b < 9}} \pi^{B_{2,1}}_{32a+b} \Big), \tag{5.31}$$

the probability that second case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 2}) = x_{2,1} \cdot \pi^{B_{2,1}}_{32N_{2,1}+1} \cdot y_{2,1}, \tag{5.32}$$

where $\pi^{B_{2,1}}$ is the stationary distribution vector such that:

$$P^{B_{2,1}} \cdot \pi^{B_{2,1}} = \pi^{B_{2,1}}.$$

In total, using the same distribution vector $\pi^{B_{2,1}}$, we can now define the average input rate of the model for buffer $B_{2,1}$ as follows as the sum of both cases:

$$R_{2,1}^{\text{input}}(x_{2,1}, y_{2,1}, P^{B_{2,1}}) \stackrel{\text{def}}{=} x_{2,1} \cdot (y_{2,1} \cdot \pi_{32N_{2,1}+1}^{B_{2,1}} + \sum_{\substack{(a,b) \in \mathcal{R}^{B_{2,1}} \\ a < N_{2,1} \\ b < 9}} \pi_{32a+b}^{B_{2,1}}). \tag{5.33}$$

Output from buffer $B_{2,1}$ can only occur in the following case:

1. Buffer $B_{2,1}$ contains at least one product part, buffer $B_{2,2}$ is not starved, buffer B_3 is not blocked, machine M_2 is up and takes a product part.

If we translate this in a formula we obtain:

$$R_{2,1}^{\text{output}}(y_{2,1}, P^{B_{2,1}}) \stackrel{\text{def}}{=} y_{2,1} \cdot (\sum_{\substack{(a,b) \in \mathcal{R}^{B_{2,1}} \\ a > 0 \\ b \in \{1,9,17,25\}}} \pi_{32a+b}^{B_{2,1}}), \tag{5.34}$$

where again $\pi^{B_{2,1}}$ is the stationary distribution vector such that:

$$P^{B_{2,1}} \cdot \pi^{B_{2,1}} = \pi^{B_{2,1}}.$$

For the stationary distribution of a buffer model we always have the conservation of flow property, which means that:

$$R_{2,1}^{\text{input}}(x_{2,1}, y_{2,1}, P^{B_{2,1}}) = R_{2,1}^{\text{output}}(y_{2,1}, P^{B_{2,1}}). \tag{5.35}$$

5.7.2 The Average Input and Output Rate of Buffer $B_{2,2}$

Input in buffer $B_{2,2}$ can only occur in one of the following two cases:

1. Buffer $B_{2,2}$ contains less than $N_{2,2}$ product parts, buffer $B_{1,2}$ is not starved, machine $M_{1,2}$ is up and produces a (non-scrapped) product part.
2. Buffer $B_{2,2}$ contains exactly $N_{2,2}$ product parts, buffer $B_{1,2}$ is not starved, machine $M_{1,2}$ is up and produces a (non-scrapped) product part *and* buffer $B_{2,1}$ is not starved, buffer B_3 is not blocked, machine M_2 is up and takes a product part.

The probability that the first case occurs can be formulated as follows:

$$\mathbb{P}\text{r}(\text{case 1}) = x_{2,2} \cdot (\sum_{\substack{(a,b) \in \mathcal{R}^{B_{2,2}} \\ a < N_{2,2} \\ b < 9}} \pi_{32a+b}^{B_{2,2}}), \tag{5.36}$$

the probability that second case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 2}) = x_{2,2} \cdot \pi^{B_{2,2}}_{32N_{2,2}+1} \cdot y_{2,2}, \tag{5.37}$$

where $\pi^{B_{2,2}}$ is the stationary distribution vector such that:

$$P^{B_{2,2}} \cdot \pi^{B_{2,2}} = \pi^{B_{2,2}}.$$

In total, using the same distribution vector $\pi^{B_{2,2}}$, we can now define the average input rate of the model for buffer $B_{2,2}$ as follows as the sum of both cases:

$$R^{\text{input}}_{2,2}(x_{2,2}, y_{2,2}, P^{B_{2,2}}) \stackrel{\text{def}}{=} x_{2,2} \cdot (y_{2,2} \cdot \pi^{B_{2,2}}_{32N_{2,2}+1} + \sum_{\substack{(a,b) \in \mathcal{R}^{B_{2,2}} \\ a < N_{2,2} \\ b < 9}} \pi^{B_{2,2}}_{32a+b}). \tag{5.38}$$

Output from buffer $B_{2,2}$ can only occur in the following case:

1. Buffer $B_{2,2}$ contains at least one product part, buffer $B_{2,2}$ is not starved, buffer B_3 is not blocked, machine M_2 is up and takes a product part.

If we translate this in a formula we obtain:

$$R^{\text{output}}_{2,2}(y_{2,2}, P^{B_{2,2}}) \stackrel{\text{def}}{=} y_{2,2} \cdot (\sum_{\substack{(a,b) \in \mathcal{R}^{B_{2,2}} \\ a > 0 \\ b \in \{1,9,17,25\}}} \pi^{B_{2,2}}_{32a+b}), \tag{5.39}$$

where again $\pi^{B_{2,2}}$ is the stationary distribution vector such that:

$$P^{B_{2,2}} \cdot \pi^{B_{2,2}} = \pi^{B_{2,2}}.$$

For the stationary distribution of a buffer model we always have the conservation of flow property, which means that:

$$R^{\text{input}}_{2,2}(x_{2,2}, y_{2,2}, P^{B_{2,2}}) = R^{\text{output}}_{2,2}(y_{2,2}, P^{B_{2,2}}). \tag{5.40}$$

5.7.3 The Average Input and Output Rate of Buffer B_3

Input in buffer B_3 can only occur in one of the following two cases:

1. Buffer B_3 contains less than N_3 product parts, both buffers $B_{1,2}$ and $B_{2,2}$ are not starved, machine M_2 is up and produces a (non-scrapped) product part.

2. Buffer B_3 contains exactly N_3 product parts, both buffers $B_{1,2}$ and $B_{2,2}$ are not starved, machine M_2 is up and produces a (non-scrapped) product part *and* buffer B_4 is not blocked, machine M_3 is up and takes a product part.

The probability that the first case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 1}) = x_3 \cdot \Big(\sum_{\substack{(a,b) \in \mathcal{R}^{B_3} \\ a < N_3 \\ b < 5}} \pi^{B_3}_{32a+b} \Big), \tag{5.41}$$

the probability that second case occurs can be formulated as follows:

$$\mathbb{P}\mathrm{r}(\text{case 2}) = x_3 \cdot \pi^{B_3}_{32N_3+1} \cdot y_3, \tag{5.42}$$

where π^{B_3} is the stationary distribution vector such that:

$$P^{B_3} \cdot \pi^{B_3} = \pi^{B_3}.$$

In total, using the same distribution vector π^{B_3}, we can now define the average input rate of the model for buffer B_3 as follows as the sum of both cases:

$$R_3^{\text{input}}(x_3, y_3, P^{B_3}) \stackrel{\text{def}}{=} x_3 \cdot \Big(y_3 \cdot \pi^{B_3}_{32N_3+1} + \sum_{\substack{(a,b) \in \mathcal{R}^{B_3} \\ a < N_3 \\ b < 5}} \pi^{B_3}_{32a+b} \Big). \tag{5.43}$$

Output from buffer B_3 can only occur in the following case:

1. Buffer B_3 contains at least one product part, buffer B_4 is not blocked, machine M_3 is up and takes a product part.

If we translate this in a formula we obtain:

$$R_3^{\text{output}}(y_3, P^{B_3}) \stackrel{\text{def}}{=} y_3 \cdot \Big(\sum_{\substack{(a,b) \in \mathcal{R}^{B_3} \\ a > 0 \\ b \in \{1,5,9,13,17,21,25,29\}}} \pi^{B_3}_{32a+b} \Big), \tag{5.44}$$

where again π^{B_3} is the stationary distribution vector such that:

$$P^{B_3} \cdot \pi^{B_3} = \pi^{B_3}.$$

For the stationary distribution of a buffer model we always have the conservation of flow property, which means that:

$$R_3^{\text{input}}(x_3, y_3, P^{B_3}) = R_3^{\text{output}}(y_3, P^{B_3}). \tag{5.45}$$

5.8 The Algorithm that Searches for Parameters that Meet the Restrictions

In order to find the parameters that satisfy the restrictions as much as possible we developed an algorithm that we will describe in this subsection. For the long term relations, the Markov transition matrices $P^{B_{2,1},\text{in}}$ and $P^{B_{2,2},\text{in}}$ of the input sub-models for buffers $B_{2,1}$ and $B_{2,2}$, as well as the Markov transition matrix $P^{B_3,\text{out}}$ of the output sub-model for buffer B_3 are supposed to be given boundary conditions. For the short term relations the input rates $x_{2,1}$ and $x_{2,2}$ of buffers $B_{2,1}$ and $B_{2,2}$, as well as the output rate y_3 are also supposed to be given boundary conditions.

The problem that the next algorithm tries to solve is: which are the long term Markov transition matrices $P^{B_{2,1},\text{out}}$, $P^{B_{2,2},\text{out}}$ and $P^{B_3,\text{out}}$ and which are the short term rates $y_{2,1}$, $y_{2,2}$ and x_3 such that both, the short term and long term restrictions are satisfied best?

5.8.1 A Detailed Description of the Algorithm

Step 1.
Initialization of the three 4×4 sized Markov transition matrices

$$P^{B_{2,1},\text{in}}(s) := \begin{pmatrix} 1-\alpha_{1,1} & \beta_{1,1} & 1-\alpha_{1,1} & \beta_{1,1} \\ \alpha_{1,1} & 1-\beta_{1,1} & \alpha_{1,1} & 1-\beta_{1,1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{for all } s \in I\!N,$$

$$P^{B_{2,2},\text{in}}(s) := \begin{pmatrix} 1-\alpha_{1,2} & \beta_{1,2} & 1-\alpha_{1,2} & \beta_{1,2} \\ \alpha_{1,2} & 1-\beta_{1,2} & \alpha_{1,2} & 1-\beta_{1,2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{for all } s \in I\!N,$$

$$P^{B_3,\text{out}}(s) := \begin{pmatrix} 1-\alpha_3 & \beta_{1,1} & 1-\alpha_3 & \beta_3 \\ \alpha_3 & 1-\beta_3 & \alpha_3 & 1-\beta_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{for all } s \in I\!N,$$

corresponding to the input sub-model for buffer $B_{2,1}$, the input sub-model for buffer $B_{2,2}$ and the output sub-model for buffer B_3.

Initialization, using equalities from (5.30), of the three scalars

$$\begin{aligned} x_{2,1}(s) &:= p_{1,1} \cdot (1 - c_{1,1}), \quad \text{for all } s \in I\!N, \\ x_{2,2}(s) &:= p_{1,2} \cdot (1 - c_{1,2}), \quad \text{for all } s \in I\!N, \\ y_3(s) &:= p_3, \quad \text{for all } s \in I\!N. \end{aligned}$$

All these transition matrices and scalars remain unchanged during the whole algorithm and operate as boundary conditions.

Step 2.
Initialization of the 8×8 sized Markov transition matrices

$$P^{B_{2,1},\text{out}}(0) := \begin{pmatrix} 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 \\ \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$P^{B_{2,2},\text{out}}(0) := \begin{pmatrix} 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 \\ \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$P^{B_3,\text{in}}(0) := \begin{pmatrix} 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 & 1-\alpha_2 & \beta_2 \\ \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 & \alpha_2 & 1-\beta_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

corresponding to the output sub-model for buffer $B_{2,1}$, the output sub-model for buffer $B_{2,2}$ and the input sub-model for buffer B_3.

Initialization of the three scalars:

$$y_{2,1}(0) := p_2 ,$$

$$y_{2,2}(0) := p_2 ,$$

$$x_3(0) := p_2 \cdot (1 - c_2) .$$

These transition matrices and scalars change during the algorithm until a satisfactory convergence is reached.

Step 3.
Initialization of the number s of the iteration step $s := 0$.

Step 4.
Construction of transition matrix $P^{B_{2,1}}(s+1)$, using the appropriate equality from (5.26):

$$P^{B_{2,1}}(s+1) := P^{B_{2,1}}(x_{2,1}(s+1), y_{2,1}(s), P^{B_{2,1},\text{in}}(s+1), P^{B_{2,1},\text{out}}(s)),$$

of the model for buffer $B_{2,1}$. An illustration for this step is shown in Figure 5.15.

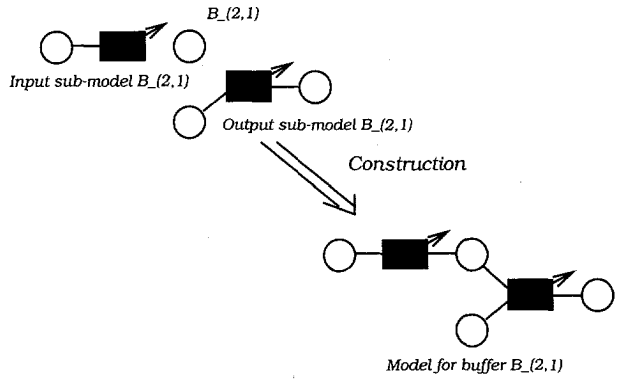

Figure 5.15: Step 4 of the algorithm.

Step 5.
Reduction in two steps of the transition matrix $P^{B_{2,1}}(s+1)$ of the model for buffer $B_{2,1}$ to the transition matrix $P^{B_3,\text{in}}(s+1)$ of the input sub-model for buffer B_3, using the appropriate equality from (5.27):

$$P_{2,1}^{\text{overlap}}(s+1) := \tilde{P}^{B_{2,1}}_{\rho_{2,1}^{\text{overlap}}}(s+1) ,$$
$$P^{B_3,\text{in}}(s+1) := \left(\tilde{P_{2,1}^{\text{overlap}}}(s+1)\right)_{\rho_{\text{overlap}}^{B_3,\text{in}}} ,$$

and computation of the new parameter $x_3(s+1)$ that determines the input rate for the number of product parts per time slot that enters buffer B_3, using the appropriate equality from (5.30):

$$x_3(s+1) := p_2 \cdot (1-c_2) \cdot (1 - \Psi_{\text{st}}^{B_{2,1}}(P^{B_{2,1}}(s+1)) \cdot (1 - \Psi_{\text{st}}^{B_{2,2}}(P^{B_{2,2}}(s)) .$$

An illustration for this step is shown in Figure 5.16.

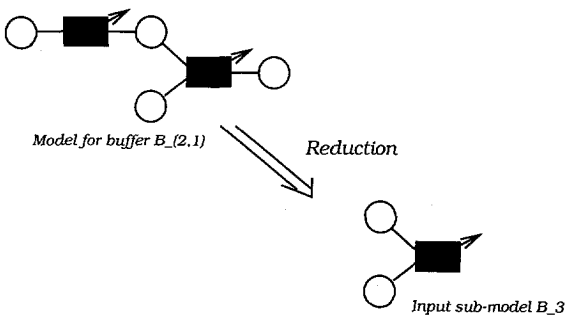

Figure 5.16: Step 5 of the algorithm.

Step 6.
Construction of transition matrix $P^{B_3}(s+1)$, using the appropriate equality from (5.26):

$$P^{B_3}(s+1) := P^{B_3}(x_3(s+1), y_3(s+1), P^{B_3,\text{in}}(s+1), P^{B_3,\text{out}}(s+1)),$$

of the model for buffer B_3. An illustration for this step is shown in Figure 5.17.

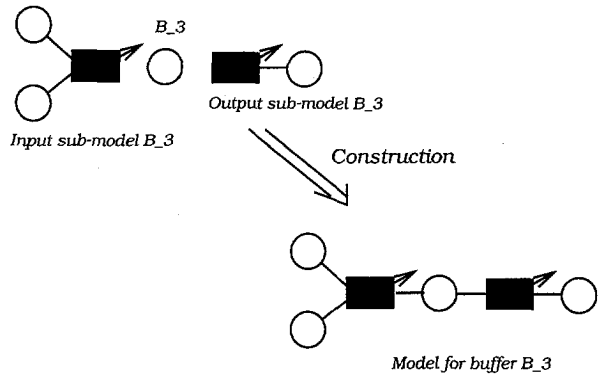

Figure 5.17: Step 6 of the algorithm.

Step 7.
Reduction in two steps of the transition matrix $P^{B_3}(s+1)$ of the model for buffer B_3 to the transition matrix $P^{B_{2,2},\text{out}}(s+1)$ of the output sub-model for buffer $B_{2,2}$, using the appropriate equality from (5.27):

$$\begin{aligned} P_3^{\text{overlap}}(s+1) &:= \tilde{P}^{B_3}_{\rho_3^{\text{overlap}}}(s+1)\ , \\ P^{B_{2,2},\text{out}}(s+1) &:= \left(\tilde{P_3^{\text{overlap}}}(s+1)\right)_{\rho_{\text{overlap}}^{B_{2,2},\text{out}}}\ . \end{aligned}$$

and computation of the new parameter $y_{2,2}(s+1)$ that determines the output rate for the number of product parts per time slot that leaves buffer $B_{2,2}$, using the appropriate equality from (5.30):

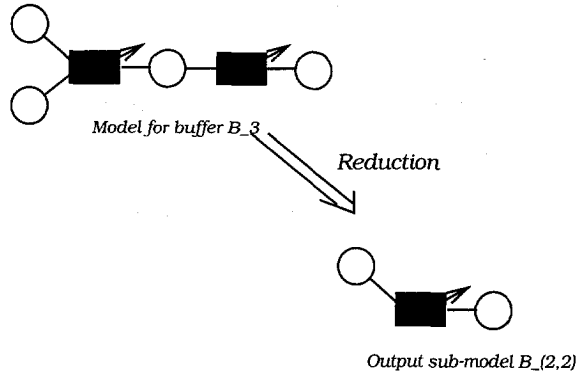

Figure 5.18: Step 7 of the algorithm.

$$y_{2,2}(s+1) := p_2 \cdot (1 - \Psi_{\text{st}}^{B_{2,1}}(P^{B_{2,1}}(s+1)) \cdot (1 - \Psi_{\text{bl}}^{B_3}(q_3, P^{B_3}(s+1)) \ ,$$

and the computation of the average production rate of the production line. The average production rate of the production line is equal to the fraction of the average output rate of buffer B_3 as described in Expression (5.44) that remains after scrapping of machine M_3:

$$R_3(s+1) := (1 - c_3) \cdot R_3^{\text{output}}(y_3(s+1), P^{B_3}(s+1)).$$

An illustration for this step is shown in Figure 5.18.

Step 8.
Construction of transition matrix $P^{B_{2,2}}(s+1)$, using the appropriate equality from (5.26):

$$P^{B_{2,2}}(s+1) \ := \ P^{B_{2,2}}(x_{2,2}(s+1), y_{2,2}(s+1), P^{B_{2,2},\text{in}}(s+1), P^{B_{2,2},\text{out}}(s+1)),$$

of the model for buffer $B_{2,2}$. An illustration for this step is shown in Figure 5.19.

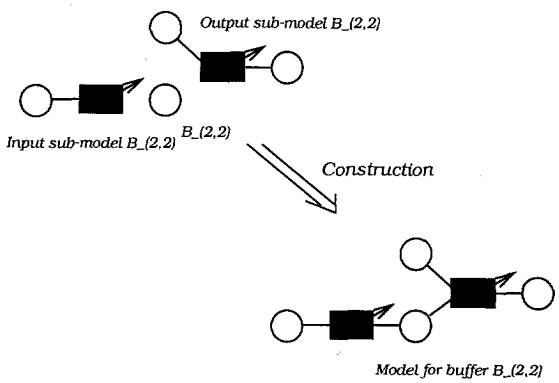

Figure 5.19: Step 8 of the algorithm.

Step 9.
Reduction in two steps of the transition matrix $P^{B_{2,2}}(s+1)$ of the model for buffer $B_{2,2}$ to the transition matrix $\bar{P}^{B_3,\text{in}}(s+1)$ of the input sub-model for buffer B_3, using the appropriate equality from (5.27):

$$\begin{aligned} P_{2,2}^{\text{overlap}}(s+1) \ &:= \ \tilde{P}^{B_{2,2}}_{\rho_{2,2}^{\text{overlap}}}(s+1) \ , \\ \bar{P}^{B_3,\text{in}}(s+1) \ &:= \ \left(\tilde{P_{2,2}^{\text{overlap}}}(s+1)\right)_{\rho_{\text{overlap}}^{B_3,\text{in}}} \ , \end{aligned}$$

and computation of the new parameter $\bar{x}_3(s+1)$ that determines the input rate for the number of product parts per time slot that enters buffer B_3, using the appropriate equality from (5.30):

$$\bar{x}_3(s+1) := p_2 \cdot (1-c_2) \cdot (1-\Psi_{\mathrm{st}}^{B_{2,2}}(P^{B_{2,1}}(s+1)) \cdot (1-\Psi_{\mathrm{st}}^{B_{2,1}}(P^{B_{2,2}}(s+1)) \ .$$

An illustration for this step is shown in Figure 5.20.

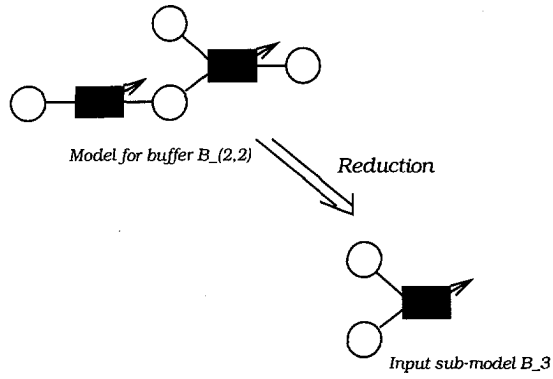

Figure 5.20: Step 9 of the algorithm.

Step 10.
Construction of transition matrix $P^{B_3}(s+1)$, using the appropriate equality from (5.26):

$$\bar{P}^{B_3}(s+1) \ := \ P^{B_3}(\bar{x}_3(s+1), y_3(s+1), \bar{P}^{B_3,\mathrm{in}}(s+1), P^{B_3,\mathrm{out}}(s+1)),$$

of the model for buffer B_3. An illustration for this step is shown in Figure 5.21.

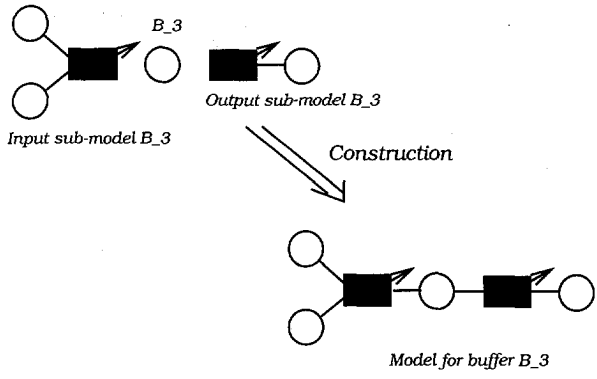

Figure 5.21: Step 10 of the algorithm.

Step 11.
Reduction in two steps of the transition matrix $\bar{P}^{B_3}(s+1)$ of the model for buffer B_3 to the transition matrix $P^{B_{2,1},\text{out}}(s+1)$ of the output sub-model for buffer $B_{2,1}$, using the appropriate equality from (5.27):

$$\begin{aligned} \bar{P}_3^{\text{overlap}}(s+1) &:= \tilde{\bar{P}}^{B_3}_{\rho_3^{\text{overlap}}}(s+1) \ , \\ P^{B_{2,1},\text{out}}(s+1) &:= \left(\tilde{\bar{P}}_3^{\text{overlap}}(s+1)\right)_{\rho_{\text{overlap}}^{B_{2,1},\text{out}}} . \end{aligned}$$

and computation of the new parameter $y_{2,1}(s+1)$ that determines the output rate for the number of product parts per time slot that leaves buffer $B_{2,1}$, using the appropriate equality from (5.30):

$$y_{2,1}(s+1) := p_2 \cdot (1 - \Psi_{\text{st}}^{B_{2,2}}(P^{B_{2,2}}(s+1)) \cdot (1 - \Psi_{\text{bl}}^{B_3}(1 - y_3(s), \bar{P}^{B_3}(s+1)) \ .$$

and the computation of the average production rate of the production line. The average production rate of the production line is equal to the fraction of the average output rate of buffer B_3 as described in Expression (5.44) that remains after scrapping of machine M_3:

$$\bar{R}_3(s+1) := (1 - c_3) \cdot R_3^{\text{output}}(y_3(s+1), \bar{P}^{B_3}(s+1)).$$

An illustration for this step is shown in Figure 5.22.

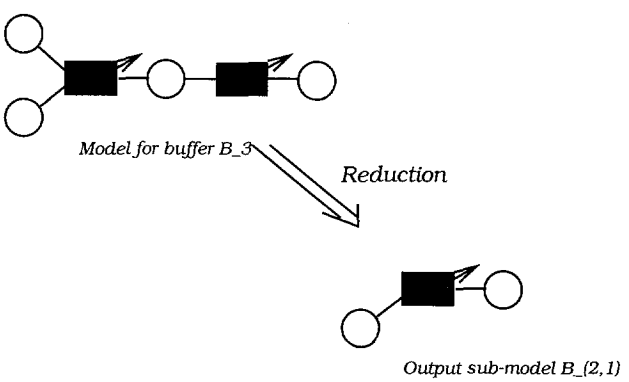

Figure 5.22: Step 11 of the algorithm.

Step 12.
If no satisfactory convergence is reached yet then $s := s + 1$ goto **Step 4**.

Step 13.
Stop.

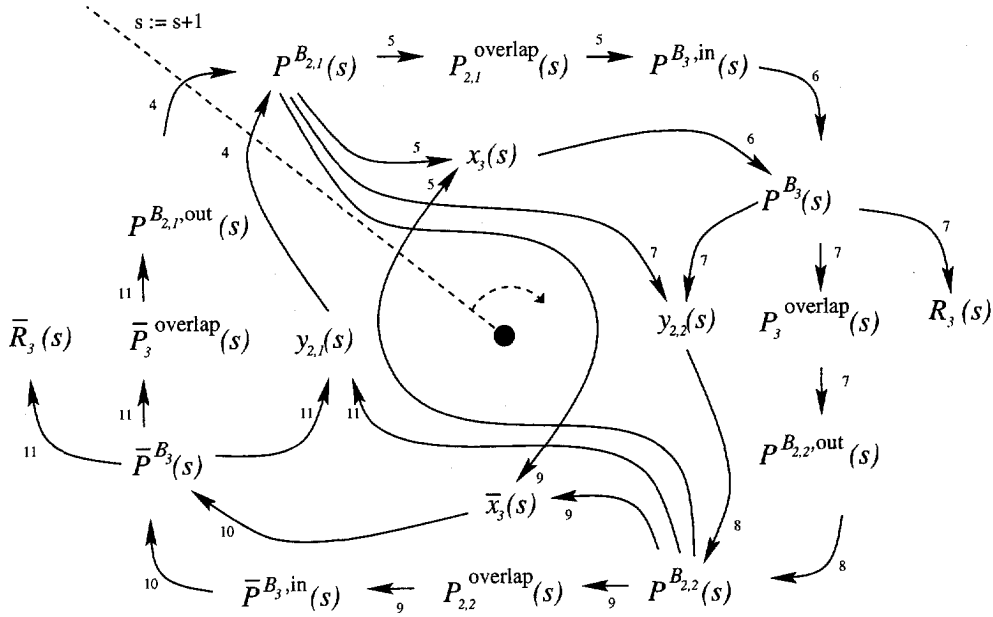

Figure 5.23: The algorithm schematically.

5.8.2 Some Comments on the Algorithm

A more schematic illustration of the algorithm is shown in Figure 5.23. Each arrow in the illustration links two elements. The relation between these elements is such that the element at the beginning of the arrow is needed, in a way that is described in the previous subsection, in order to compute the element at the end of the arrow. Crossing the dashed line implies that the number s of the iteration cycle increases by one. We can see in this illustration that within one iteration cycle s we obtain four separate estimates, $P_{2,1}^{\text{overlap}}(s)$, $P_3^{\text{overlap}}(s)$, $P_{2,2}^{\text{overlap}}(s)$, $\bar{P}_3^{\text{overlap}}(s)$ for the Markov transition matrix of the model for the overlap. As stated in Expression (5.25) we would like, as the iteration cycle number s increases, that these estimates of the transition matrix of the model for physically the same object will converge to the same matrix as well. However, this is *not* the case in general. It turns out that, in all cases that we have examined, the algorithm converges numerically in the sense that the following limits appear to exist:

$$\lim_{s\to\infty} P_{2,1}^{\text{overlap}}(s) \stackrel{\text{def}}{=} P_{2,1}^{\text{overlap}}, \quad \lim_{s\to\infty} P_{2,2}^{\text{overlap}}(s) \stackrel{\text{def}}{=} P_{2,2}^{\text{overlap}},$$

$$\lim_{s\to\infty} P_3^{\text{overlap}}(s) \stackrel{\text{def}}{=} P_3^{\text{overlap}}, \quad \lim_{s\to\infty} \bar{P}_3^{\text{overlap}}(s) \stackrel{\text{def}}{=} \bar{P}_3^{\text{overlap}},$$

and

$$\lim_{s\to\infty} y_{2,1}(s) \stackrel{\text{def}}{=} y_{2,1} \ , \quad \lim_{s\to\infty} y_{2,2}(s) \stackrel{\text{def}}{=} y_{2,2} \ ,$$

$$\lim_{s\to\infty} x_3(s) \stackrel{\text{def}}{=} x_3 \ , \quad \lim_{s\to\infty} \bar{x}_3(s) \stackrel{\text{def}}{=} \bar{x}_3 \ .$$

We have to accept that the Markov transition matrices for the overlap model generated by means of the algorithm do not converge to equality. Therefore we cannot find the matrices $P_{2,1}^{\text{overlap}}$, P_3^{overlap}, $P_{2,2}^{\text{overlap}}$ and $\bar{P}_3^{\text{overlap}}$ that satisfy Expression (5.25) by means of this algorithm:

$$P_{2,1}^{\text{overlap}} \neq P_{2,2}^{\text{overlap}} \neq P_3^{\text{overlap}} \neq \bar{P}_3^{\text{overlap}}. \tag{5.46}$$

Apart from that, the algorithm computes two different scalars x_3 and $\bar{x}_3$ which both correspond to the number of product parts per time slot that enter buffer B_3. The algorithm does not guarantee either that these scalars end up equal:

$$x_3 \neq \bar{x}_3. \tag{5.47}$$

The fact that we do not find the equal transition matrices for the model of the overlap makes us suspect that the solution to the combination of all restrictions in (5.25), (5.26), (5.27) and (5.30) does not exist. The fact that we cannot find equal transition matrices is probably due to the way of construction of the transition matrices $P^{B_{2,1}}$, $P^{B_{2,2}}$ and P^{B_3} based on *independent* input and output sub-processes. The input and output sub-processes of each of the buffers are in fact not independent. The fit of our buffer models, and with that the success of the algorithm is therefore totally dependent on the level of interdependence of the input and output sub-models for each of the buffers.

The fact that we do not find equal transition matrices does not at all mean that our algorithm is useless since, fortunately, in the cases that we have examined the differences between the matrices $P_{2,1}^{\text{overlap}}$, P_3^{overlap}, $P_{2,2}^{\text{overlap}}$ and $\bar{P}_3^{\text{overlap}}$, as well as the difference between x_3 and $\bar{x}_3$, these differences remained very small.

5.9 Application of the Algorithm to the Car Lamp Production Line

In order to apply our algorithm to the car lamp production line we have to find the best fitting parameters for the breakdown and repair rates and for the speeds and scrap probabilities of all the machines that are involved in the production line.

5.9.1 The Identification Procedure

From three machines in the production line we have been able to collect detailed data about events that occurred in the car lamp production line during some hours. However, it

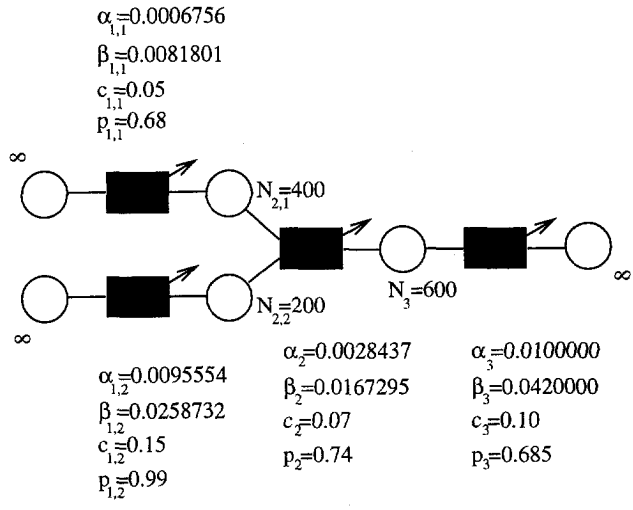

Figure 5.24: The parameters that we use in this chapter for the car lamp production line.

turned out that these data were not sufficient to perform a proper identification procedure for the parameters involved. Several problems had to be tackled in a poor identification procedure. According to the data, for instance, machines went down when they were in down state already or went up when they were already in up state. Next to that the time of recording was too short to draw conclusions and determine breakdown and repair rates. Another problem was the recording of the data we used which was not performed in parallel for the separate machines as we wished. From the machines we received data from totally different epochs in time. In spite of all these problems we tried our best to make our model fit best with the processes in the car lamp production line. Parameters that could have come from our identification procedure are shown in Figure 5.24[1]. For the production line with these parameters we applied the algorithm that we have described before in full detail.

5.9.2 Application of the Algorithm

The first thing that is important to examine is the convergence of the algorithm as the iteration cycles increase. This convergence in the sense that we described in Expressions (5.46) and (5.46) was examined by taking the norms of differences between corresponding objects from subsequent iteration cycles. The results for the four different Markov tran-

[1]The values in Figure 5.24 are *not* the values from the identification procedure. For reasons of industrial secrecy we have changed the values.

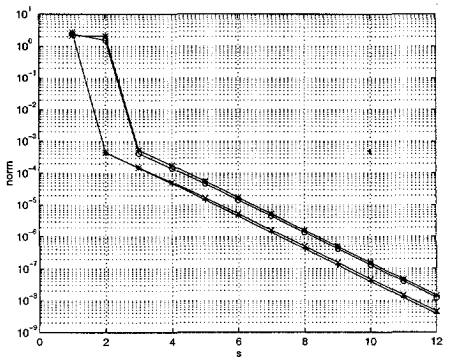

Figure 5.25: $\| X(s) - X(s-1) \|$ for $X = P_{2,1}^{\text{overlap}}$, P_3^{overlap}, $P_{2,2}^{\text{overlap}}$ and $\bar{P}_3^{\text{overlap}}$ we used "*","o","x" and "+" respectively.

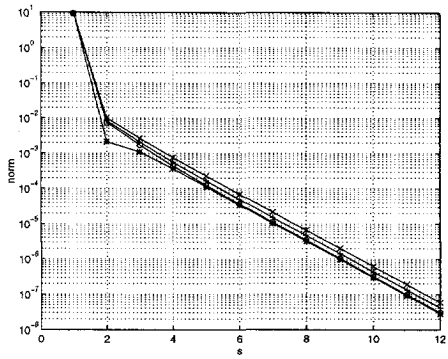

Figure 5.26: $| x(s) - x(s-1) |$ for $x = y_{2,1}$, x_3, $y_{2,2}$ and $\bar{x}_3$ we used "*","o","x" and "+" respectively.

sition matrices for the model of the overlap are shown in Figure 5.25. The results for the four different input or output rates are shown in Figure 5.26. From the fact that the subsequent differences persist in lying on a straight descending line we draw the conclusion that convergence is attained for all the transition matrices and input and output rates involved in the algorithm.

The result, after 12 iteration cycles of the algorithm, of scalars $y_{2,1}(12)$, $x_3(12)$, $y_{2,2}$, $\bar{x}_3(12)$ matrices $P_{2,1}^{\text{overlap}}(12)$, $P_3^{\text{overlap}}(12)$, $P_{2,2}^{\text{overlap}}(12)$ and $\bar{P}_3^{\text{overlap}}(12)$ are listed in Appendix C in Tables C.11, C.12, C.13, C.14 and C.15 respectively. For each of these 4 Markov transition matrices of the overlap we computed also the stationary distributions which are shown in the last row of each of the tables. From the stationary distribution we can compute, for each of the transition matrices for the overlap, some probabilities for blocking and starvation of the buffers. The results from each of the transition matrix of the overlap is shown in Figure 5.2.

The next thing that we examine is the differences between the four transition matrices for the overlap that come from the algorithm. We explained already that the algorithm does not guarantee the desired equalities in Expression (5.25). In order for our results to make sense we need that the differences between the matrices are relatively small. A table of distances between the different Markov transition matrices for physically the same overlap is shown in Table 5.3. In this table we can see that the norms of the differences between the different Markov transition matrices are all less than 0.003. The fact that these differences fortunately are very small gives us the strength to continue with our examinations.

	$P_{2,1}^{\text{overlap}}(12)$	$P_{3}^{\text{overlap}}(12)$	$P_{2,2}^{\text{overlap}}(12)$	$\bar{P}_{3}^{\text{overlap}}(12)$
$\mathbb{P}\text{r}(B_{2,1}^{\neg\text{st}})$	0.9847	0.9847	0.9847	0.9847
$\mathbb{P}\text{r}(B_{2,1}^{\text{st}})$	0.0153	0.0153	0.0153	0.0153
$\mathbb{P}\text{r}(B_{2,2}^{\neg\text{st}})$	0.9583	0.9583	0.9583	0.9583
$\mathbb{P}\text{r}(B_{2,2}^{\text{st}})$	0.0417	0.0417	0.0417	0.0417
$\mathbb{P}\text{r}(B_{3}^{\neg\text{bl}})$	0.9938	0.9938	0.9938	0.9940
$\mathbb{P}\text{r}(B_{3}^{\text{bl}})$	0.0062	0.0062	0.0062	0.0060
$\mathbb{P}\text{r}(M_{2}^{\text{up}})$	0.8547	0.8547	0.8547	0.8547
$\mathbb{P}\text{r}(M_{2}^{\text{down}})$	0.1453	0.1453	0.1453	0.1453

Table 5.2: The states for the separate objects in the overlap and their computed probability.

$X \backslash Y$	$P_{2,1}^{\text{overlap}}(12)$	$P_{3}^{\text{overlap}}(12)$	$P_{2,2}^{\text{overlap}}(12)$	$\bar{P}_{3}^{\text{overlap}}(12)$
$P_{2,1}^{\text{overlap}}(12)$	0	0.0005	0.0021	0.0021
$P_{3}^{\text{overlap}}(12)$	0.0005	0	0.0019	0.0019
$P_{2,2}^{\text{overlap}}(12)$	0.0021	0.0019	0	0.0018
$\bar{P}_{3}^{\text{overlap}}(12)$	0.0021	0.0019	0.0018	0

Table 5.3: A distance table. For combination X and Y we used distance $\| X - Y \|_2$. $|x_3(12) - \bar{x}_3(12)| \approx 1.5\ 10^{-9}$.

5.9.3 Computation Results Compared with Simulation Results

The computed stationary distributions $\pi^{B_{2,1}}$, $\pi^{B_{2,1}}$ and π^{B_3} will be represented by means of the distributions of the contents of the three buffers. In order to evaluate the results from the algorithm we performed also simulations of the car lamp production line with the parameters as shown in Figure 5.24. An important tool for such simulations is the random generator used. For our simulations in C we used again the *drand48* command which we have used also in examples from Chapter 2, Chapter 3 and Chapter 4 in Section 2.8, Section 3.11 and Section 4.6 respectively. The simulations are performed here in a similar way as we performed the simulations in the corresponding sections in Chapter 2, Chapter 3 and Chapter 4. During a simulation of 10 million time slots we kept track of the relative frequencies of the contents of each buffer. We also derived the steady state distribution of the contents of each buffer as a result from the computations by means of the algorithm. In order to compare the simulation results with the results from the algorithm, for each buffer, we have put the computed stationary distribution of the contents next to the estimated stationary distribution by means of relative frequencies from the simulation. The results for the stationary distribution of the contents of buffer $B_{2,1}$ are shown in Figures 5.27 and 5.28. The results for the stationary distribution of the contents of buffer $B_{2,2}$ are shown in Figures 5.29 and 5.30. The results for the stationary distribution of the contents of buffer B_3 are shown in Figures 5.31 and 5.32.

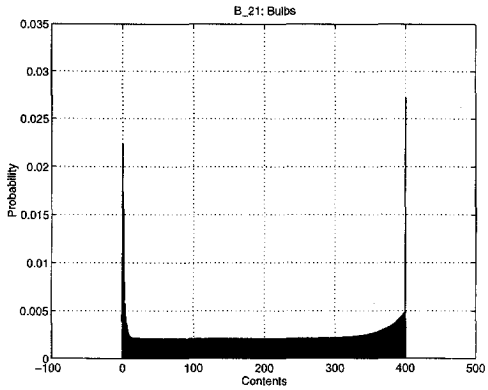

Figure 5.27: The *computed* stationary distribution of the contents of buffer $B_{2,1}$.

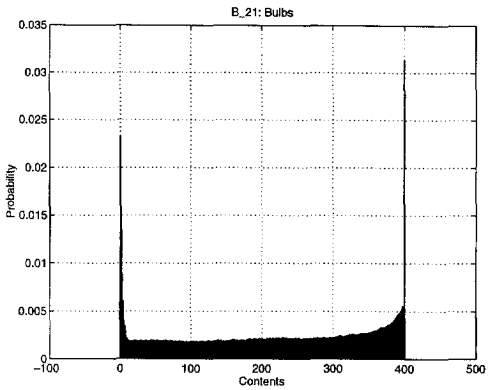

Figure 5.28: The *simulated* stationary distribution of the contents of buffer $B_{2,1}$.

Actually the computed distribution of the contents of buffer B_3 in Figure 5.31 is one of the *two* distributions for this buffer computed by means of the algorithm. The algorithm produces after 12 iteration cycles (s=12) *two* Markov transition matrices of the model for buffer B_3: the model $P^{B_3}(12)$ as constructed in the last time that **Step 6** was executed, and the model $\bar{P}^{B_3}(12)$ as constructed in the last time that **Step 10** was executed. Both transition matrices are not equal, and neither are the corresponding stationary distribution vectors for the contents of the buffer. The stationary distribution vector in Figure 5.31 is based on the stationary distribution of the last transition matrix $\bar{P}^{B_3}(12)$. The stationary distribution vector for the contents of buffer B_3 based on the matrix $P^{B_3}(12)$ is similar to the stationary distribution vector shown in Figure 5.31. The differences between the two stationary distributions are invisible on the same scale since the norm of the difference between both distribution vectors is less than 5.10^{-4}. The difference between the two distribution vectors is shown in Figure 5.33. We subtracted the stationary distribution with $\bar{P}^{B_3}(12)$ from the stationary distribution with $P^{B_3}(12)$. The fact that we obtain two different Markov transition matrices for buffer B_3 has the consequence that we will obtain also two different average production rates $R_3(12)$ and $\bar{R}_3(12)$ of the car lamp production line from the algorithm, each of them based on one of the two models.

Next to the simulation for the stationary distribution of the buffer contents we also performed 10 independent runs of simulations of one million time slots each. Before we did these 10 independent runs we first performed one initial run of one million time slots in order to estimate the average contents of each buffer. Each of the following 10 independent simulation runs start in a position with average buffer contents so as to minimize the influence that the transient behaviour has on the estimates of the stationary distribution.

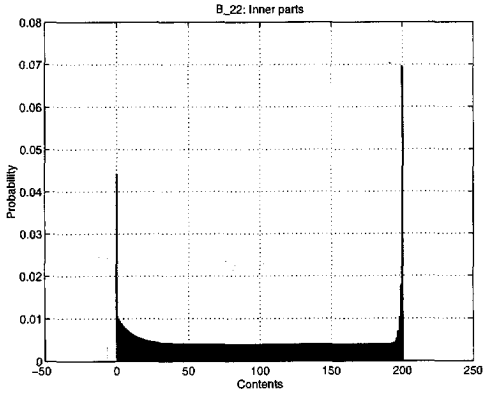

Figure 5.29: The *computed* stationary distribution of the contents of buffer $B_{2,2}$.

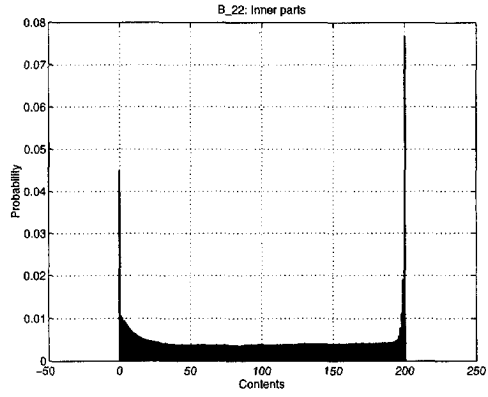

Figure 5.30: The *simulated* stationary distribution of the contents of buffer $B_{2,2}$.

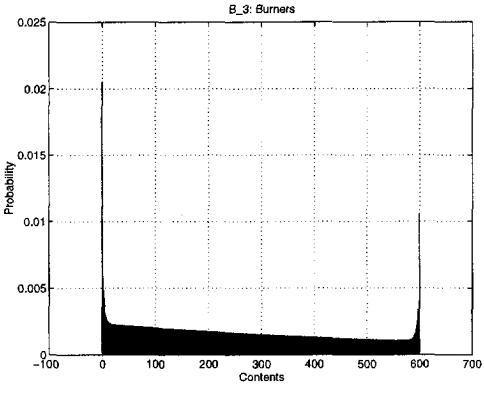

Figure 5.31: The *computed* stationary distribution of the contents of buffer B_3.

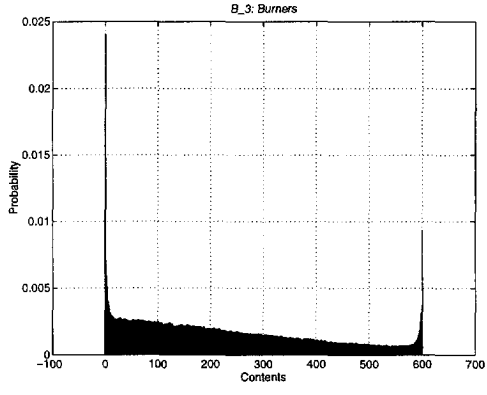

Figure 5.32: The *simulated* stationary distribution of the contents of buffer B_3.

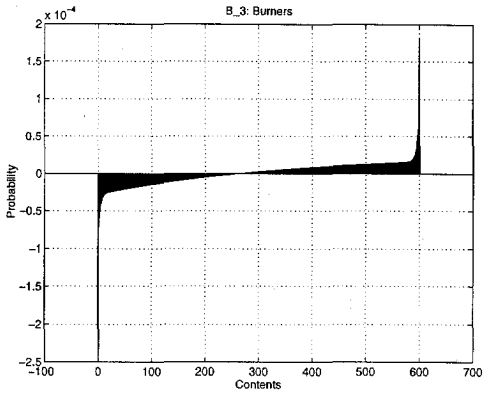

Figure 5.33: The difference between the two computed stationary distribution vectors of the contents of buffer B_3.

Estimates of average production rate R by simulation										
	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
R	.485872	.486399	.483776	.484725	.487023	.485598	.485463	.486402	.483572	.484293

Estimates of the average buffer contents by simulation										
Buf.	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
B_2	227.644	223.791	198.044	186.975	248.812	210.532	213.515	276.495	263.490	201.926
B_3	222.564	192.972	207.458	208.946	218.099	230.123	218.592	227.462	199.608	209.647
B_4	98.611	100.060	103.241	100.897	101.037	96.458	100.969	99.289	109.221	102.080

Table 5.4: The results of 10 runs of simulation for the car lamp production line simulating 10^6 time steps each

By means of these 10 independent runs we computed 10 independent average production rates and for each of the three buffers 10 independent numbers for the average contents. By means of each of these corresponding 10 independent numbers we can calculate the mean and the standard deviation of the average production rate and the average buffer contents. The results from the 10 independent runs of simulation are shown in Table 5.4. The means and standard deviations, compared with the mean values that come from our algorithm, of the average production rate of the car lamp production line and the average contents of each buffer are shown in Table 5.5.

5.9.4 Sensitivity Analysis for the Car Lamp Production Line

It is nice that we now have an algorithm, described in Section 5.8, by means of which, given the breakdown and repair rates, machine speeds, scrap probabilities and buffer capacities

Results for the average production rate R						
	mean value	std. deviation	decomp. model		abs. err.	rel. err.
R	0.4853123	0.0011778	$R_3(12)$	0.4871643	0.0018520	0.0038016
			$\bar{R}_3(12)$	0.4870366	0.0017243	0.0035395

Results for the average buffer contents						
Buffer	mean value	std. deviation	decomp. model		abs. err.	rel. err.
$B_{2,1}$	213.5	12.0	206.4		- 7.1	-0.033
$B_{2,2}$	101.2	3.4	99.3		- 1.9	-0.019
B_3	225.1	29.4	1	259.2	34.1	0.151
			2	257.7	32.6	0.145

Table 5.5: The compound results of the simulations for the car lamp production line

we can compute an average production rate of the production line as a whole. But, for purposes in practice, it is even more interesting to know what happens if we change the values of these variables by *addition or subtraction of small numbers.* If we know how small changes of the values of variables in the production line effect the overall production rate then we have a measure for how "critical" these variables are. In such a way we obtain a measure for the importance of those variables for the performance of the production line in question. Such a measure for the importance of a variable in the model for the overall performance of the production line will be denoted by the term *sensitivity.*

In order to find these sensitivities we simply changed each of the values of the 19 variables in the model, one at a time, by addition and substraction of small numbers. Then we recompute with each new set of values the average production rate by means of the algorithm that we described in Section 5.8.

Note that by means of the algorithm in Section 5.8 from each set of values we obtain *two* estimates of the average production rate denoted by R_3 and $\bar{R}_3$, which ideally are identical. The fact that they are not identical implies that we also obtain *two* sets of sensitivities, each of which corresponds to either R_3 or $\bar{R}_3$. Within each of these two sets we destinguish between two kinds of sensitivies, the *absolute sensitivity* and the *relative sensitivity.* Absolute sensitivity here denotes the *actual* increase of the average production rate divided by the *actual* increase of the value of the variable in question. Relative sensitivity here denotes the *percentage* of increase of the average production rate divided by the *percentage* of increase of the value of the variable in question. The results of the sensitivity analysis for the long term variables, the short term variables and the buffer capacities are shown in

Var.	Base	R_3 0.4871643		Sensitivity		$\bar{R}_3$ 0.4870366		Sensitivity	
	Value	+10%	−10%	Abs.	Rel. 10^{-2}	+10%	−10%	Abs.	Rel.
$\alpha_{1,1}$	.0006756	.4856931	.4883913	−20.0	−2.77	.4855283	.4882954	−20.5	−2.84
$\alpha_{1,2}$	.0095554	.4817892	.4904740	−4.54	−8.91	.4816156	.4903966	−4.59	−9.01
α_2	.0028437	.4844819	.4895572	−8.92	−5.21	.4843648	.4894243	−8.90	−5.19
α_3	.0100000	.4843582	.4882123	−1.92	−3.96	.4843092	.4880223	−1.86	−3.81
$\beta_{1,1}$	.0081801	.4885983	.4850601	2.16	3.63	.4885138	.4848667	2.23	3.74
$\beta_{1,2}$	.0258732	.4907356	.4801535	2.04	10.86	.4906577	.4799847	2.06	10.96
β_2	.0167295	.4900520	.4831509	2.06	7.08	.4899256	.4830297	2.06	7.08
β_3	.0420000	.4882013	.4837146	0.53	4.60	.4880133	.4836725	0.516	4.45

Table 5.6: Sensitivities of the *long term* variables: the breakdown and repair rates.

Var.	Base	R_3 0.4871643		Sensitivity		$\bar{R}_3$ 0.4870366		Sensitivity	
	Value	+1%	−1%	Abs.	Rel.	+1%	−1%	Abs.	Rel.
$p_{1,1}$	.680	.4882692	.4856943	.189	.264	.4881664	.4855365	.193	.270
$p_{1,2}$	.990	.4883986	.4856437	.139	.283	.4882872	.4854994	.141	.286
p_2	.740	.4883056	.4858834	.164	.249	.4881704	.4857650	.163	.247
p_3	.685	.4878158	.4860615	.128	.180	.4876528	.4859755	.122	.172
$c_{1,1}$	.050	.4870968	.4872308	−.134	−.014	.4869677	.4871045	−.137	−.014
$c_{1,2}$	.150	.4869181	.4874015	−.161	−.050	.4867875	.4872768	−.163	−.050
c_2	.070	.4868614	.4874645	−.431	−.062	.4867308	.4873398	−.435	−.063
c_3	.100	.4866230	.4877056	−.541	−.111	.4864954	.4875777	−.541	−.111

Table 5.7: Sensitivities of the *short term* variables: the speeds and scrap probabilities.

Table 5.6, Table 5.7 and Table 5.8 respectively.

5.10 Conclusions

In this chapter we tried successfully to adapt and apply the methods developed in previous chapters for serial production lines to a non-serial production line in a car lamp factory with merging.

We decomposed the car lamp production line in three big Markov chain models, of which each corresponds to a single buffer and of which each consists of several Markov chain sub-models. The long term relations between the Markov chain models were dealt with by means of reduction to clusters that correspond to states of an overlap between the models.

Variable	Base	R_3 0.4871643		Sensitivity		$\bar{R}_3$ 0.4870366		Sensitivity	
	Value	+10%	−10%	Abs.	Rel. 10^{-3}	+10%	−10%	Abs.	Rel.
$N_{2,1}$	400	.4877289	.4864855	1.55	12.8	.4876153	.4863414	1.59	13.1
$N_{2,2}$	200	.4882321	.4858318	6.00	24.6	.4881187	.4856889	6.07	24.9
N_3	600	.4874186	.4868406	0.48	5.9	.4872827	.4867219	0.47	5.8

Table 5.8: Sensitivities of the buffer capacities.

The short term relations were dealt with by means of compensation of the speed for the rates in which the buffers are short term blocked or starved. In this way we found restrictions for the connections between the Markov chain models for the three buffers.

On the basis of these restrictions we developed an iterative algorithm that searches for unknown transition matrices and rates such that most of the restrictions are satisfied. We applied this algorithm to the car lamp production line and conclude that it does converge in the following sense. The algorithm does not converge to a single limit, but it ends up switching alternately between *two* limits that are very close to each other. This implies that the solution from the algorithm does not satisfy all the restrictions. Therefore we think that, because of approximate assumptions on independency of input and output sub-processes, in general a solution that meats all the restrictions does not exist. Fortunately, for the application in the car lamp production line the two solutions from the algorithm are very close. They both meet most restrictions and, because they are close, they meet *almost* the rest of the restrictions. A measure for the way in which the solution meets restrictions is shown in a distance table.

As a consequence of the special way of convergence the algorithm finds *two* different estimates, R_3 and $\bar{R}_3$, of the average production rate. These two estimates correspond to *two* stationary distributions that are very close to each other.

We also compared the computed stationary distribution with simulations and conclude that the results from computations are accurate. Only one estimate was slightly out of range, this was the estimate of the average contents of buffer B_3.

Next to the computation of the average production rate we did computations in order to estimate the sensitivity to small changes in the values of the parameters in the model. The sensitivities that came from the computations are listed in tables and can be used to make decisions for changes in the car lamp production line.

Chapter 6

An Approach with Timed Event Graphs

In the previous chapters we have tried to estimate the average production rate of production lines in which machines go up and down by means of a simple breakdown-repair model with parameters called breakdown and repair rate. In up state the service times of machines were geometrically distributed with a parameter called completion probability. This means that the service times of the machines (including the up and down transitions) in the models from previous chapters have very specific properties, such that they can easily be described by means of Markov chains. In this chapter we will introduce a totally new description of serial production lines with blocking using *stochastic timed event graphs* which are a subclass of the stochastic timed Petri nets. This new approach allows us to introduce service times as well as transportation times with *arbitrary* distributions. Before we describe serial production lines by means of stochastic timed event graphs we will first explain some elementary properties of timed event graphs and how they work. For a more thorough description of Petri nets and stochastic timed event graphs we refer to [BCOQ92].

6.1 Introduction to Stochastic Timed Event Graphs

6.1.1 Description of a Timed Event Graph

A timed event graph is a graph that consists of *places*, *transitions*, *arcs*, *holding times* and *tokens*. Symbols for these items are shown in Figure 6.1. The properties of such a graph are as follows:

- Arcs connect the places and transitions. Arcs can only be directed from transitions to places or from places to transitions. There are no arcs from transitions to transitions or from places to places.

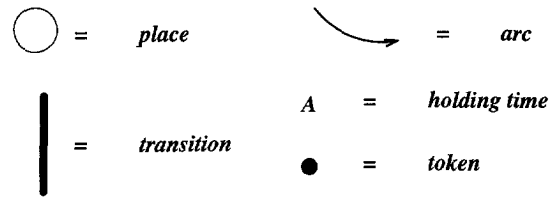

Figure 6.1: The parts of a timed event graph

- Each transition has *upstream places* and *downstream places*. An upstream place is a place from which an arc is directed to the transition. A downstream place is a place to which an arc is directed from the transition.

- Each place has *upstream transitions* and *downstream transitions*. An upstream transition is a transition from which an arc is directed to the place. A downstream transition is a transition to which an arc is directed from the place.

- Places have at most one upstream and at most one downstream transition. (This distinguishes an event graph from a Petri net. For Petri nets this property should be dropped.)

- Every place has its own holding time which can be either stochastic or deterministic.

- Transitions or arcs do not have holding times.

- A token travels through the graph with special rules for traveling. By means of the rules tokens can "split up" and "join", which means that the total number of tokens in a timed event graph is not a constant in time. Tokens travel in such a way that at every possible moment every token belongs to a single place.

- The rules for the traveling of tokens are such that they keep repeating the following procedure. This *prevents tokens from overtaking*:

  - Enter a place.
  - Wait until all previous tokens in the same place have left and until additional conditions for firing of the next transition are satisfied.
  - Leave the place.

- Transitions can *fire* as soon as all upstream places have at least one token that has waited longer than a prescribed holding time. If a transition fires one token leaves from each of its upstream places and one token enters in each of its downstream places.

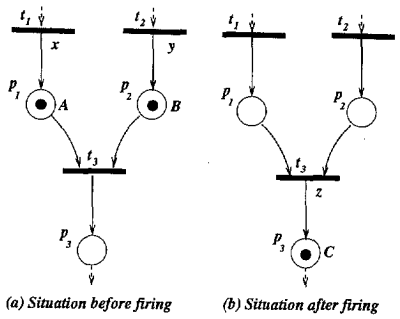

Figure 6.2: Firing of a transition

An example is shown in Figure 6.2.

Example.
Figure 6.2 shows two similar timed event graphs with three transitions t_1, t_2, t_3 and three places p_1, p_2, p_3. Let us focus on transition t_3. It has two upstream places, namely places p_1 and p_2, and it has only one downstream place, place p_3. In the following we explain a firing of this transition. Suppose that transition t_1 fired at time x and transition t_2 fired at time y, then at time x one token was put into place p_1 and at time y one token was put into place p_2. This situation is shown in Figure 6.2 (a). The token in place p_1 has to wait there for A time units, since the holding time of place p_1 is A, whereas the token in place p_2 has to wait there B time units, since the holding time of place p_2 is B. The next transition t_3 fires as soon as both tokens have finished waiting. Transition t_3 fires at time z. Now it is easy to derive the relation between x, y and z: $z = \max(x + A, y + B)$. When transition t_3 fires one token is subtracted from all upstream places, places p_1 and p_2, and one token is added to all downstream places, place p_3. The situation after firing of transition t_3 is shown in Figure 6.2 (b). There is only one token that arrives in place t_3 after firing, since there is only one downstream place. This token arrives in place p_3 at time z and has to wait there C time units, since place p_3 has holding time C... etcetera.

□

6.1.2 A Simple Autonomous Stochastic Timed Event Graph

The subject of this subsection is a simple timed event graph with only two transitions, t_1 and t_2 and four places, p_1, p_2, p_3 and p_4 as shown in Figure 6.3. The holding times of the k-th token that arrives at places p_1, p_2, p_3, p_4 are A_k, B_k, C_k, D_k respectively. The index k stands for the k-th token that arrives, so A_k denotes the holding time of the k-th token that arrives at place p_1 etcetera. $x_1(k)$ denotes the time instant of the k-th firing of transition t_1, whereas $x_2(k)$ denotes the time instant of the k-th firing of transition t_2. The

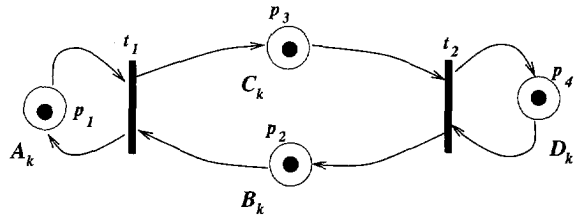

Figure 6.3: The two-dimensional timed event graph model

following assumptions are made on the stochastic holding times of the timed event graph in Figure 6.3.

- $\omega(k)$ is for all $k \in \mathbb{N}$ a random vector on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = \mathbb{R}^4$ and $\mathbb{P}$ is a probability measure on a σ-algebra $\mathcal{F}$ associated with Ω, such that $A_k = \omega_1(k)$, $B_k = \omega_2(k)$, $C_k = \omega_3(k)$, $D_k = \omega_4(k)$. For the sake of convenience we will take $\mathcal{F} = \mathcal{B}(\mathbb{R}^4)$, the collection of Borel sets on $\mathbb{R}^4$. Note that in this description A_k, B_k, C_k and D_k may be dependent.
- $\omega(i)$ is independent of $\omega(j)$ for all $i \neq j$.

6.1.3 A Mathematical Description of the Mechanism

From Figure 6.3 it is easy to derive the equations that govern the system introduced in the previous subsection. Transition t_1 can fire for the $k+1$-st time only if the k-th token in place p_1 and the k-th token in place p_2 have finished their holding times. The k-th token in place p_1 finishes its holding time after it has arrived in place p_1 at time $x_1(k)$ and then has waited for A_k time units. The k-th token in place p_2 finishes its holding time after it has arrived in place p_2 at time $x_2(k)$ and then has waited for B_k time units. All this can be put in the following formula:

$$x_1(k+1) = \max(\, x_1(k) + A_k \,,\, x_2(k) + B_k \,). \tag{6.1}$$

In a similar way we can derive the formula for the $k+1$-st firing of transition t_2:

$$x_2(k+1) = \max(\, x_1(k) + C_k \,,\, x_2(k) + D_k \,). \tag{6.2}$$

Equations (6.1) and (6.2) completely define the mechanism for the simple stochastic timed event graph. We will assume that the initial conditions in the timed event graph are such that exactly at time 0 both transitions t_1 and t_2 fire for the 0^{th} time:

$$\begin{aligned} x_1(0) &= 0\,, \\ x_2(0) &= 0\,. \end{aligned}$$

6.1.4 The Evolution of Distributions of x

First we define the joint distribution function of $x_1(k)$ and $x_2(k)$:

$$F_{x(k)}(u,v) \stackrel{\text{def}}{=} \mathbb{P}\text{r}(\ x_1(k) \leq u \ \wedge \ x_2(k) \leq v\ ). \tag{6.3}$$

Then, evidently:

$$F_{x(1)}(u,v) = \begin{cases} 1, & \text{for } u \geq 0 \ \wedge \ v \geq 0 \\ 0 & \text{elsewhere} \end{cases}. \tag{6.4}$$

It is not difficult to derive an iteration of joint distribution functions of $x_1(k)$ and $x_2(k)$ by means of the formulas (6.1) and (6.2):

$$\begin{aligned} F_{x(k+1)}(u,v) &= \mathbb{P}\text{r}(x_1(k+1) \leq u \ \wedge \ x_2(k+1) \leq v), \\ &= \mathbb{P}\text{r}(\max(x_1(k)+A_k, x_2(k)+B_k) \leq u \ \wedge \\ &\qquad \max(x_1(k)+C_k, x_2(k)+D_k) \leq v), \\ &= \mathbb{P}\text{r}(x_1(k) \leq \min(u-A_k, v-C_k) \ \wedge \\ &\qquad x_2(k) \leq \min(u-B_k, v-D_k)), \\ &= \int_{\Omega} F_{x(k)}(\min(u-A_k, v-C_k), \min(u-B_k, v-D_k))\ \mathrm{d}\mathbb{P}. \end{aligned} \tag{6.5}$$

Given the initial joint distribution $F_{x(k)}$, the distribution function $F_{x(k+1)}$ is formally fixed by Equation (6.5). It is possible to numerically perform the integration step in Equation (6.5) and iterate for $k = 1, 2, 3, ...$, but that would not give much insight in the throughput of the system. In order to obtain more information about this behavior, we use a slightly different approach in the following subsection.

6.1.5 A Partial De-Coupling of the System

To de-couple the system partially we perform the following linear transformation and introduce the variables $y_1(k)$ and $y_2(k)$:

$$\begin{pmatrix} y_1(k) \\ y_2(k) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix}. \tag{6.6}$$

Now we can combine (6.1), (6.2) and (6.6) and get the following mechanism:

$$\begin{aligned} y_1(k+1) &= y_1(k) + \tfrac{1}{2}\{\max(A_k + y_2(k), B_k - y_2(k)) + \\ &\qquad \max(C_k + y_2(k), D_k - y_2(k))\}, \\ y_2(k+1) &= \tfrac{1}{2}\{\max(A_k + y_2(k), B_k - y_2(k)) - \\ &\qquad \max(C_k + y_2(k), D_k - y_2(k))\} \end{aligned} \tag{6.7}$$

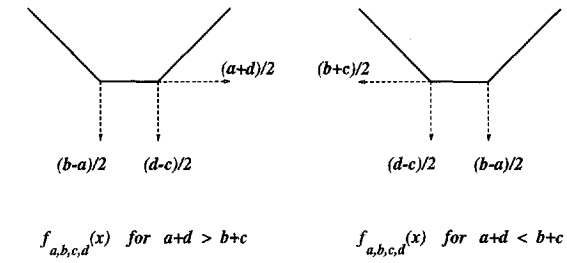

Figure 6.4: The function $f_{a,b,c,d}(x)$

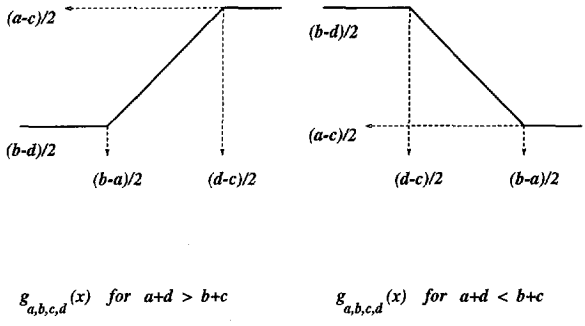

Figure 6.5: The function $g_{a,b,c,d}(x)$

Although mechanism (6.7) looks more complicated, it will give more insight in the limiting behaviour of the system. Notice that we managed to de-couple the evolution of $y_2(k)$ from the evolution of $y_1(k)$. The evolution of $y_2(k)$ is autonomous and does not depend on $y_1(k)$. The evolution of $y_1(k)$ however is not independent of $y_2(k)$, therefore we call it a partial de-coupling. Now we define the following functions:

$$\begin{aligned} f_{a,b,c,d}(x) &\stackrel{\text{def}}{=} \tfrac{1}{2}\{\max(a+x, b-x) + \max(c+x, d-x)\}, \\ g_{a,b,c,d}(x) &\stackrel{\text{def}}{=} \tfrac{1}{2}\{\max(a+x, b-x) - \max(c+x, d-x)\}. \end{aligned} \tag{6.8}$$

Sketches of these functions $f_{a,b,c,d}$ and $g_{a,b,c,d}$ are shown in Figures 6.4 and 6.5. By means of the definitions in (6.8) we can reformulate Equation (6.7) as follows:

$$\begin{aligned} y_1(k+1) &= y_1(k) + f_{A_k,B_k,C_k,D_k}(y_2(k)) \\ y_2(k+1) &= g_{A_k,B_k,C_k,D_k}(y_2(k)) \end{aligned} \tag{6.9}$$

From (6.9) follows that

$$y_1(k) \;=\; \sum_{i=1}^{k} f_{A_i,B_i,C_i,D_i}(y_2(i)). \tag{6.10}$$

6.1.6 A Coordinate Transformation in the Sample Space Ω

For a more convenient notation we transform the sample space Ω by means of the linear coordinate transformation defined by matrix M:

$$\bar{\omega}(k) \;\stackrel{\text{def}}{=}\; M\,\omega(k),$$

where

$$M \;=\; \frac{1}{4}\begin{pmatrix} -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

$$M^{-1} \;=\; 4M'.$$

With this transformation in the sample space we will define 4 new random variables:

$$\begin{aligned} X_k &\stackrel{\text{def}}{=} \bar{\omega}_1(k), \\ Y_k &\stackrel{\text{def}}{=} \bar{\omega}_2(k), \\ \Gamma_k &\stackrel{\text{def}}{=} \bar{\omega}_3(k), \\ \Lambda_k &\stackrel{\text{def}}{=} \bar{\omega}_4(k). \end{aligned}$$

By means of the new random variables X_k, Y_k, Γ_k and Λ_k we can simplify the evolution of $y_1(k)$ and $y_2(k)$ from Expression (6.9):

$$\begin{aligned} y_1(k+1) &= y_1(k) + \Lambda_k + \frac{|(y_2(k) - X_k) + \Gamma_k| + |(y_2(k) - X_k) - \Gamma_k|}{2}, \\ y_2(k+1) &= Y_k + \frac{|(y_2(k) - X_k) + \Gamma_k| - |(y_2(k) - X_k) - \Gamma_k|}{2}. \end{aligned} \tag{6.11}$$

If we define a new state variable $\bar{y}_2(k)$ as follows:

$$\bar{y}_2(k) \;\stackrel{\text{def}}{=}\; y_2(k) - X_k,$$

and we define a new random variable

$$Z_k \;\stackrel{\text{def}}{=}\; Y_k - X_k,$$

then we can reformulate the evolution of $y_1(k)$ and $y_2(k)$ in the most convenient way:

$$\begin{aligned} y_1(k+1) &= y_1(k) + \Lambda_k + \max(\,|\Gamma_k|\,,\,|\bar{y}_2(k)|\,) \\ \bar{y}_2(k+1) &= \frac{|\bar{y}_2(k) + \Gamma_k| - |\bar{y}_2(k) - \Gamma_k|}{2} + Z_k. \end{aligned} \tag{6.12}$$

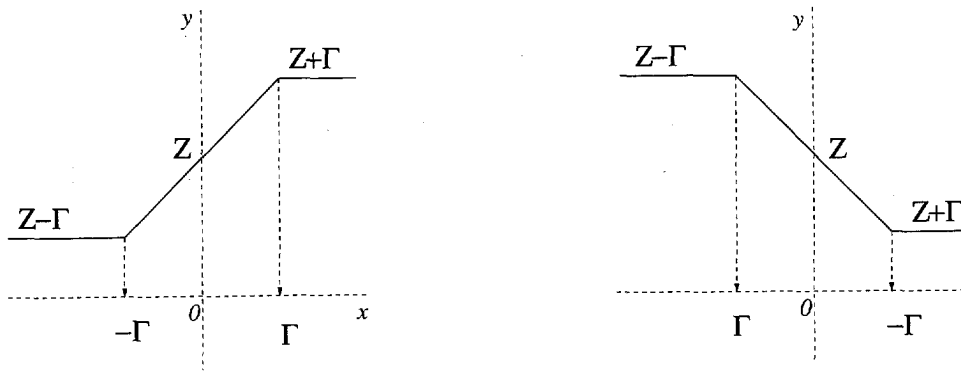

Figure 6.6: The function $y = \Phi_{\Gamma,Z}(x)$.

If we define, for $\gamma,\ z \in \mathbb{R}$, the kernel function $\Phi_{\gamma,z} : \mathbb{R} \mapsto \mathbb{R}$ as follows:

$$\Phi_{\gamma,z}(x) \stackrel{\text{def}}{=} \frac{|x+\gamma| - |x-\gamma|}{2} + z, \tag{6.13}$$

then we obtain the following for Expression (6.12):

$$\begin{aligned} y_1(k+1) &= y_1(k) + \Lambda_k + \max(\ |\Gamma_k|\ ,\ |\bar{y}_2(k)|\ ) \\ \bar{y}_2(k+1) &= \Phi_{\Gamma_k,Z_k}(\ \bar{y}_2(k)\ ). \end{aligned} \tag{6.14}$$

A picture of the function $y = \Phi_{\Gamma,Z}(x)$ is shown in Figure 6.6.

Remarks

Expression (6.14) is an interesting result because it shows that:

- The concurrency in the system, represented by the evolution of $\bar{y}_2(k)$, is determined *only* by the two sequences $\{\Gamma_k\}$, $\Gamma_k = \frac{1}{4}(A_k - B_k - C_k + D_k)$ and $\{Z_k\}$, $Z_k = \frac{1}{2}(A_k - D_k)$. For examination of the concurrency in the original system it is best to examine the concurrency in the following equivalent system:

$$\begin{aligned} x_1(k+1) &= \max(\ x_1(k) + \Gamma_k + Z_k,\ \ x_2(k) - \Gamma_k\ ) \\ x_2(k+1) &= \max(\ x_1(k) - \Gamma_k,\ \ x_2(k) + \Gamma_k - Z_k\ ). \end{aligned} \tag{6.15}$$

- The evolutions of the systems corresponding to the sequences (A_k, B_k, C_k, D_k) and $(A_k, B_k + V_k, C_k - V_k, D_k)$, for some sequence $\{V_k\}$ with $V_k \in \mathbb{R}$, are completely equal. As a consequence, without loss of generality, we can restrict ourselves to the "symmetric" sequences $(A_k, \frac{B_k + C_k}{2}, \frac{B_k + C_k}{2}, D_k)$.

6.1.7 The Evolution of Distributions of $\bar{y}_2$

If we examine Expression (6.14) closely we see that the distribution of $\bar{y}_2(k+1)$ depends, by means of the function Φ_{Γ_k, Z_k} and the values of Γ_k and Z_k, on the previous distribution of $\bar{y}_2(k)$. As an extension of the transition matrices for homogeneous discrete-time finite Markov chains, here we come up with the *transition kernel* for homogeneous discrete time *infinite* Markov chains. Such a transition kernel can transform distribution functions on an interval into other distribution functions on another interval. This operation can not be described by means of matrix multiplication, but can be described by means of an integral notation. The transition kernel here is a two-dimensional function, and will be denoted by $K(s,t)$. Properties of a transition kernel $K(s,t)$ are as follows:

$$\begin{aligned} K(s,t) &\leq K(s,u), \quad \text{for all } s, \text{for all } t \leq u, \\ \lim_{t \to \infty} K(s,t) &= 1, \quad \text{for all } s \in \mathbb{R}, \\ \lim_{t \to -\infty} K(s,t) &= 0, \quad \text{for all } s \in \mathbb{R}. \end{aligned}$$

The integral notation for the transformation, corresponding to one time step, of a distribution function π into a distribution function ν is as follows:

$$\nu(t) = \int_{\mathbb{R}} K(s,t) \; \mathrm{d}\pi(s).$$

In order to find the transition kernel $K(s,t)$ for the transition from $\bar{y}_2(k)$ to $\bar{y}_2(k+1)$ we examine, for $(s,t) \in \mathbb{R}^2$, the set $W(s,t) \subset \mathbb{R}^2$ such that

$$W(s,t) \stackrel{\text{def}}{=} \{ (\gamma, z) \in \mathbb{R}^2 \mid \Phi_{\gamma,z}(s) \leq t \}. \tag{6.16}$$

It is easy to see with the definition of function $\Phi_{\gamma,z}(s)$ that then

$$W(s,t) = \{ (\gamma, z) \in \mathbb{R}^2 \mid z \leq \Phi_{-s,t}(\gamma) \}.$$

By means of these sets we can construct the kernel $K(s,t)$ of the system:

$$K(s,t) \stackrel{\text{def}}{=} \int_{\Omega} I_{W(s,t)}(\; (\Gamma_k(\omega), Z_k(\omega)) \;) \; \mathrm{d} \; \mathbb{P}.$$

By means of this kernel we derive the evolution of the distribution functions $F_{\bar{y}_2(k)}$ defined as:

$$F_{\bar{y}_2(k)}(s) \stackrel{\text{def}}{=} \mathbb{P}\text{r}\ (\ \bar{y}_2(k)\ \leq\ s\ ),$$

as follows:

$$F_{\bar{y}_2(k+1)}(t)\ =\ \int_{\mathbb{R}} K(s,t)\ \mathrm{d}F_{\bar{y}_2(k)}(s). \tag{6.17}$$

6.1.8 The Case for $Z_k = 0$ Almost Always for all k

In case that $Z_k = 0$ almost always for all k, which is the case if the measure of the subspace with $\omega_1 = \omega_4$ of the sample space Ω equals 1,

$$\int_{\Omega} I_{\{\omega_1(k)=\omega_4(k)\}}(\omega(k))\ \mathrm{d}I\!P\ \ =\ \ 1.$$

In this case we can construct a distribution function $\Gamma(x)$ as follows:

$$\Gamma(x)\ \stackrel{\text{def}}{=}\ \int_{\Omega} I_{\{\omega_1(k)+\omega_4(k)-\omega_2(k)-\omega_3(k)\leq 4x\}}(\omega(k))\ \mathrm{d}I\!P.$$

If we now construct the set $W(s,t)$ excluding sets of measure 0 we get:

$$W(s,t)\ =\ \begin{cases} \emptyset, & \text{for } t < -|s|, \\ \{\ (x,0)\ |\ x \in \mathbb{R}\ \}, & \text{for } t \geq\ |s|, \\ \{\ (x,0)\ |\ x \in [-t,\ \rightarrow)\ \}, & \text{for } s \leq -|t|\ \wedge\ s \neq -t, \\ \{\ (x,0)\ |\ x \in (\ \leftarrow,\ t]\ \}, & \text{for } s \geq\ |t|\ \wedge\ s \neq\ t. \end{cases}$$

Taking the measure of these sets leads to the kernel $K(s,t)$:

$$K(s,t)\ =\ \begin{cases} 0, & \text{for } t < -|s|, \\ 1, & \text{for } t \geq\ |s|, \\ 1 - \lim\limits_{\delta \downarrow 0} \Gamma(-t-\delta), & \text{for } s \leq -|t|\ \wedge\ s \neq -t, \\ \Gamma(t), & \text{for } s \geq\ |t|\ \wedge\ s \neq\ t. \end{cases}$$

If we substitute this in Expression (6.17) then we obtain:

$$F_{\bar{y}_2(k+1)}(t)\ =\ F_{\bar{y}_2(k)}(t) + \Gamma(t) - \lim_{\delta \downarrow 0} \left[F_{\bar{y}_2(k)}(t)\ \Gamma(|t+\delta|) + F_{\bar{y}_2(k)}(-t-\delta)\ \Gamma(-|t+\delta|)\right], \tag{6.18}$$

which, if we assume that $F_{\bar{y}_2(k)}(t)$ and $\Gamma(t)$ are continuous, transforms to:

$$F_{\bar{y}_2(k+1)}(t)\ =\ F_{\bar{y}_2(k)}(t) + \Gamma(t) - \left[F_{\bar{y}_2(k)}(t)\ \Gamma(|t|) + F_{\bar{y}_2(k)}(-t)\ \Gamma(-|t|)\right]. \tag{6.19}$$

If we define the distribution functions of the absolute values of $\bar{y}_2(k)$ and Γ_k as follows:

$$F_{|\bar{y}_2(k)|}(t) \stackrel{\text{def}}{=} \begin{cases} F_{\bar{y}_2(k)}(t) - \lim_{\delta \downarrow 0} F_{\bar{y}_2(k)}(-t-\delta), & \text{for } t \geq 0, \\ 0, & \text{for } t < 0, \end{cases}$$

$$|\Gamma|(t) \stackrel{\text{def}}{=} \begin{cases} \Gamma(t) - \lim_{\delta \downarrow 0} \Gamma(-t-\delta), & \text{for } t \geq 0, \\ 0, & \text{for } t < 0, \end{cases},$$

Then the continuous case in (6.19) can be transformed in:

$$F_{|\bar{y}_2(k+1)|}(t) = \{1 - |\Gamma|(t)\} \; F_{|\bar{y}_2(k)|}(t) + |\Gamma|(t). \tag{6.20}$$

Now we define $|\Gamma|_{\min}$, the minimum value of $|\Gamma_k|$ as follows:

$$|\Gamma|_{\min} \stackrel{\text{def}}{=} \inf\{\, t \in I\!R \mid \; |\Gamma|(t) > 0 \,\}.$$

By means of $|\Gamma|_{\min}$ we can deduce how the distribution of $|\bar{y}_2(k)|$ evolves in a pointwise limit:

$$\lim_{k \to \infty} F_{|\bar{y}_2(k)|}(t) = \begin{cases} 1, & \text{for } t > |\Gamma|_{\min}, \\ F_{|\bar{y}_2(1)|}(t), & \text{for } t < |\Gamma|_{\min}, \end{cases} \tag{6.21}$$

For $t = |\Gamma|_{\min}$ we deduce:

$$\lim_{k \to \infty} F_{|\bar{y}_2(k)|}(|\Gamma|_{\min}) = \begin{cases} 1 & \text{if } \; |\Gamma|(|\Gamma|_{\min}) > 0, \\ F_{|\bar{y}_2(1)|}(|\Gamma|_{\min}) & \text{if } \; |\Gamma|(|\Gamma|_{\min}) = 0, \end{cases} \tag{6.22}$$

From Expression (6.22) we can conclude that in the limit the following holds:

$$\lim_{k \to \infty} \; I\!Pr(\; |\bar{y}_2(k)| \; \leq \; |\Gamma_k| \;) \quad = \quad 1, \tag{6.23}$$

Or equivalently:

$$\lim_{k \to \infty} \; I\!Pr(\; \max(\; |\Gamma_k|, \; |\bar{y}_2(k)| \;) = |\Gamma_k| \;) \quad = \quad 1. \tag{6.24}$$

If we combine this with Expression (6.14) we obtain:

$$\lim_{k \to \infty} \; I\!Pr(\; y_1(k+1) - y_1(k) \; = \; \Lambda_k + |\Gamma_k| \;) \quad = \quad 1. \tag{6.25}$$

The last expression tells us that, for $A_k = D_k$, the properties of the time between two subsequent firings of each transition are determined by the the random variable

$$y_1(k+1) - y_1(k) \quad = \quad \Lambda_k + |\Gamma_k| \quad = \quad \frac{\max\{\; A_k + D_k, \; B_k + C_k \;\}}{2}. \tag{6.26}$$

Expression (6.14) shows also that if $A_k \neq D_k$ then this random variable is a lower bound for the time time between two firings.

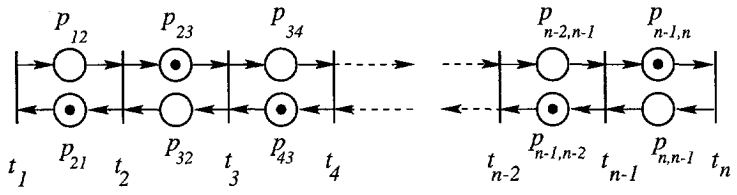

Figure 6.7: The timed event graph of a serial production line with blocking.

Figure 6.8: The two states of a single *position* with number i.

6.2 The Stochastic Timed Event Graph of a Serial Production Line with Blocking

The most general timed event graph of a serial production line with a "blocking after service" policy is shown in Figure 6.7. In this figure we can see a timed event graph with $n \in \mathbb{N}^+$ transitions t_i with $i \in \{1, 2, 3, ..., n\}$, and $2(n-1)$ places $p_{i,j}$ with $i, j \in \{1, 2, 3, ..., n\}$ and $|i - j| = 1$. The subscripts i, j of a place $p_{i,j}$ denote that the place takes tokens from upstream transition t_i and feeds tokens to a downstream transition t_j. The timed event graph in Figure 6.7 consists of several identical smaller event graphs that we will denote by means of the term *position*. Such a basic smaller event graph is shown in Figure 6.8.

In Figure 6.8 we can see clearly that, for any $i \in \{1, 2, 3, ..., n-1\}$ at *any* time instant either place $p_{i,i+1}$ has one token and $p_{i+1,i}$ has none, or place $p_{i+1,i}$ has one token and $p_{i,i+1}$ has none. It is impossible for places $p_{i,i+1}$ and $p_{i+1,i}$ to have a token at the same time instant. It is also impossible for each place to have more than one token. If a position with number i has a token in place $p_{i,i+1}$ then we will say that the position is *occupied*. If a position with number i has a token in place $p_{i+1,i}$ then we will say that the position is *free*.

Furthermore we will assume that:

- the k-th token that arrives at a place $p_{i,j}$, $|i-j|=1$, has a stochastic holding time denoted by $T_{i,j}(k)$. $T_{i,j}(k)$, for each k, is a sample from the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_{i,j})$, where $\mathbb{R}$ is the sample space, $\mathcal{B}(\mathbb{R})$ is the collection of Borel sets in $\mathbb{R}$, and $\mathbb{P}_{i,j}$ is a corresponding probability measure.

- $T_{i,j}(m)$ is independent of $T_{k,l}(n)$ for $n \neq m$ or $(i,j) \neq (k,l)$.

- $x_i(k)$, $i \in \{1,2,...,n\}$, $k \in \mathbb{N}$ denotes the time instant of the firing of transition number i for the k-th time.

- the initial conditions are as follows:
  $x_i(0) = 0$ for all even $i \in \{1,2,3,...,n\}$, $x_i(0) = -\infty$ for all odd $i \in \{1,2,3,...,n\}$.
  (This corresponds to the token distribution from Figure 6.7 in the case that n is odd.)

6.2.1 Examples: Timed Event Graphs of Production Lines from Previous Chapters

In order to show that the stochastic timed event graph indeed is of the most general form we will give some examples of serial production lines from previous chapters and describe them by means of a similar stochastic timed event graph model.

Example: the two machines production lines

In Figure 6.9 we can see a stochastic timed event graph that is equivalent to a two machines production line with a buffer between the machines that has capacity $n-2$. The two machines production line is equivalent to the "basic" two machine production lines that we considered in Chapters 2 up until 4.

- Since we assumed in previous chapters that there is no transportation time inside the buffers, we will assume that the probability measure $\mathbb{P}_{i,j}$ is such that the holding times $T_{i,j}(k) = 0$ almost always for all $k \in \mathbb{N}$, $|i-j|=1$ and $i+j \notin \{3,\ 2n-1\}$. However, this restriction on zero transportation times is not necessary in our stochastic timed event graph model. In our stochastic timed event graph model we allow any possible stochastic transportation time we want. The main restrictions on the stochastic transportation times inside the buffer are on the level of the independencies as described before.

- We will assume that the holding time $T_{1,2}(k)$ equals the service time of the k-th product part that enters machine M_1. This service time can have any possible distribution determined by the probability measure $\mathbb{P}_{1,2}$. Since we assumed in previous chapters that the machine is ready to serve the next product part just after it has finished the part before, we will assume that $T_{2,1}(k) = 0$ almost always for all $k \in \mathbb{N}$. Again, this is not a real restriction. In our timed event graph model we can allow

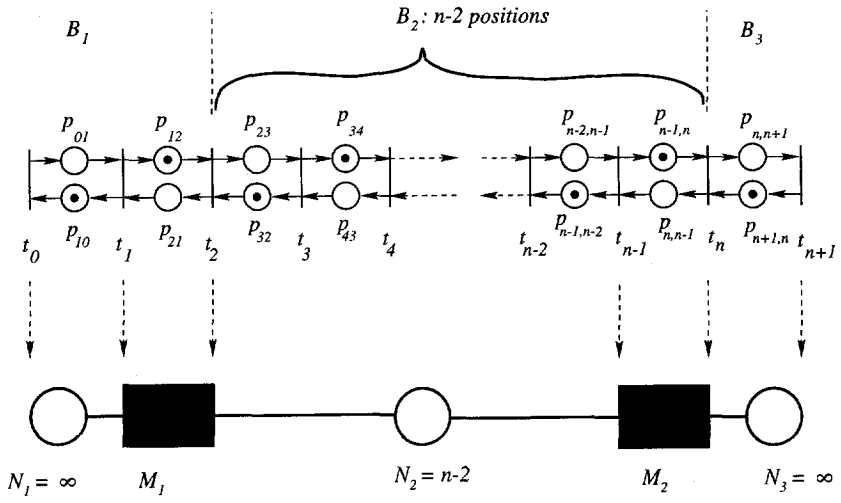

Figure 6.9: The timed event graph equivalent to a two machines production line with a buffer capacity $n-2$.

machine M_1 to need some stochastic time in order to do some preparations for each next product part.

- We will assume that the holding time $T_{n-1,n}(k)$ equals the service time of the k-th product part that enters machine M_2. This service time can have any possible distribution determined by the probability measure $\mathbb{P}_{n-1,n}$. Since we assumed in previous chapters that the machine is ready to serve the next product part just after it has finished the part before, we will assume that $T_{n,n-1}(k) = 0$ almost always for all $k \in \mathbb{N}$. Again, this is not a real restriction. In our timed event graph model we can allow machine M_2 to need some stochastic time in order to do some preparations for each next product part.

- We can see clearly in Figure 6.9 that in the Chapters 2 up until 4 we assumed that the position *inside* machine M_i is counted as one position *in* the upstream buffer B_i. (It is shown most clearly for $i = 2$.)

□

Example: the simple three machines production line from Chapter 2

In Figure 6.10 the stochastic timed event graph equivalent to the simple three machines production line from Chapter 2 is shown. Because of the buffer sizes equal to 1, which

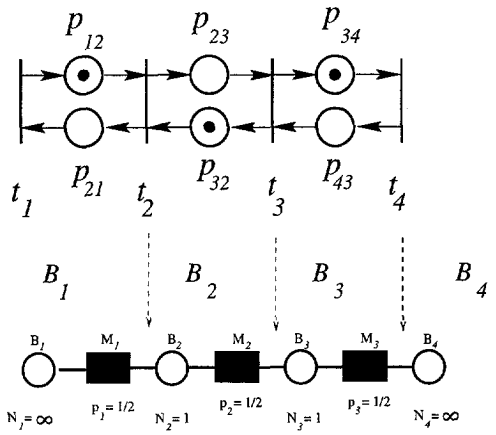

Figure 6.10: The timed event graph equivalent to the simple three machines production line from Chapter 2.

should include the position in in the following machine, there are no positions other than those inside the machines. Here, in order to obtain a production line equal to the one shown in Chapter 2, we will assume that:

- that a machine is able to start manipulation of a new product part as soon as the previous product part is finished. This implies that $I\!P_{i+1,i}$ for $i \in \{1,2,3\}$ is such that $T_{i+1,i}(k) = 0$ almost always for all $k \in I\!N$.
- that $I\!P_{i,i+1}$ for $i \in \{1,2,3\}$ is such that it corresponds to the service time distribution of machine M_i.

□

6.2.2 Mathematical Descriptions of the Evolution of the Stochastic Timed Event Graphs for Serial Production Lines

There are several different mathematical descriptions for the evolution of the stochastic timed event graph that we discuss here. The three possible descriptions that we describe in the following three subsections are denoted by the terms *odd-even*, *forward* and *backward* descriptions respectively.

The odd-even description

For the odd-even description we have several assumptions:

- We add the dummy variables defined as follows:

$$\begin{array}{ll} x_0(k) \stackrel{\text{def}}{=} -\infty, & x_{n+1}(k) \stackrel{\text{def}}{=} -\infty, \\ T_{0,1}(k) \stackrel{\text{def}}{=} -\infty, & T_{n+1,n}(k) \stackrel{\text{def}}{=} -\infty \end{array} \quad , \qquad \forall k \in I\!N, \tag{6.27}$$

  With these variables we can more easily setup the odd-even description of the stochastic timed event graph.

- The initial conditions are, as shown in Figure 6.7, such that we start with a token distribution with all odd positions free and all even positions occupied.
  All the tokens have just started their holding times. The even transitions just fired for the 0-th time: $x_i(0) = 0$ for all even $i \in \{1, 2, 3, ..., n\}$, the odd transitions have not fired $x_i(0) = -\infty$ for all odd $i \in \{1, 2, 3, ..., n\}$.

The corresponding odd-even description is as follows:

$$\begin{array}{rcll} x_i(k) & = & \max\{\ x_{i-1}(k) + T_{i-1,i}(k),\ x_{i+1}(k) + T_{i+1,i}(k)\ \}, & \text{for all odd } i, \\ x_i(k+1) & = & \max\{\ x_{i-1}(k) + T_{i-1,i}(k),\ x_{i+1}(k) + T_{i+1,i}(k)\ \}, & \text{for all even } i, \end{array} \tag{6.28}$$

where $i \in \{1, 2, 3, ..., n\}$ and $k \in I\!N$.

Example.
In order to obtain a "feeling" for what is happening in such a system we have made two illustrations shown in Figures 6.11 and 6.12. In Figure 6.11 we can see evolutions in time of four single positions in isolation, which means that each position does not wait for synchronization with neighbouring positions. We see that a token at position number i goes back and forth between transitions t_i and t_{i+1} without waiting at each of them. Figure 6.12 shows us the evolution in time for the same case but then with the synchronization obtained by means of the synchronization method in Expression (6.28). Here we can see a token at position number i go back and forth between transitions t_i and t_{i+1} *with* an appropriate waiting time at each of them. The synchronization rules in (6.28) seem simple, but the relations between the stochastic times $T_{i,j}(k)$, for $i, j \in \{1, 2, ..., n\}$, $|i - j| = 1$, $k \in I\!N$ and the waiting times for the tokens are quite complicated.

□

The forward description

In Figure 6.13 we can see the initial conditions corresponding to the forward description of timed event graph for the evolution of serial production lines. The initial conditions can be formulated as follows:

- All positions are free.

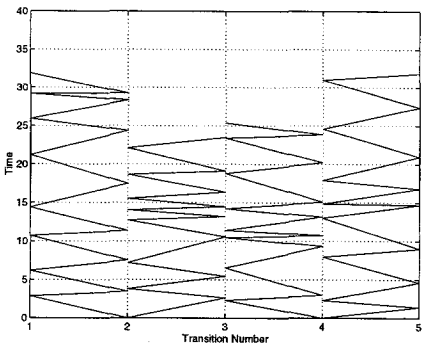

Figure 6.11: An example of a behaviour *before* synchronization.

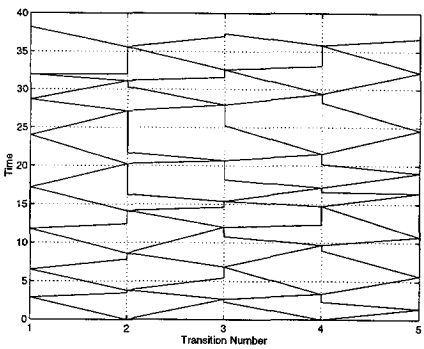

Figure 6.12: The behaviour from Figure 6.11 *after* synchronization.

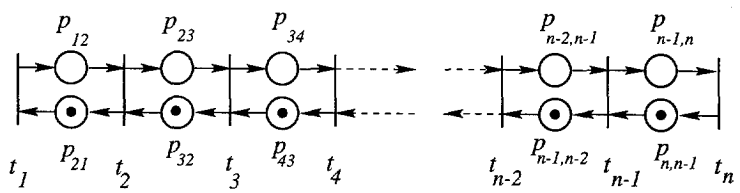

Figure 6.13: The initial token distribution corresponding to the *forward description*.

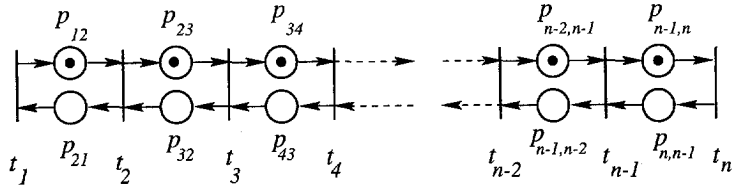

Figure 6.14: The initial token distribution corresponding to the backward description.

- All the tokens have just started their holding times. $x_i(0) = 0$ for all $i \in \{2, 3, ..., n\}$, $x_1(0) = -\infty$.

The corresponding forward description is as follows:

$$\begin{aligned} x_1(k+1) &= x_2(k) + T_{2,1}(k), \\ x_i(k+1) &= \max\{\, x_{i-1}(k+1) + T_{i-1,i}(k),\ x_{i+1}(k) + T_{i+1,i}(k)\ \}, \quad \text{for } i \in \{2, 3, ..., n\}. \end{aligned} \tag{6.29}$$

The backward description

In Figure 6.14 we can see the initial conditions corresponding to the *backward description* of timed event graph for the evolution of serial production lines. The initial conditions can be formulated as follows:

- All positions are occupied.
- All the tokens have just started their holding times. $x_i(0) = 0$ for all $i \in \{1, 2, 3, ..., n-1\}$, $x_n(0) = -\infty$.

The corresponding backward description is as follows:

$$\begin{aligned} x_n(k+1) &= x_{n-1}(k) + T_{n-1,n}(k), \\ x_i(k+1) &= \max\{\, x_{i+1}(k+1) + T_{i+1,i}(k),\ x_{i-1}(k) + T_{i-1,i}(k)\ \}, \quad \text{for } i \in \{n-1, n-2, ..., 1\}. \end{aligned} \tag{6.30}$$

6.2.3 An Approximation Method for the Distribution of Times Between Firings

We have found an appealing method in order to approximate the *distribution* of the times between two subsequent firings at a transition in the stochastic timed event graph. The distributions of times between subsequent firings of transitions determine the distribution

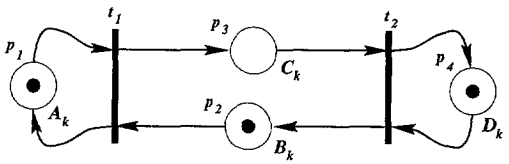

Figure 6.15: The Wait & Switch system with its initial token distribution.

of the production rate of the corresponding production line. Before we describe the method we will first examine a small and simple basic timed event graph that we will denote by the term *Wait & Switch system.*

The Wait & Switch system

The Wait & Switch system is a simple stochastic timed event graph that consists of one single position plus two "external" loops, as shown in Figure 6.15.

For this Wait & Switch system we have the following extra assumptions:

- A_k, B_k, C_k and D_k are stochastic holding times of the k-th token that arrives at places p_1, p_2, p_3 and p_4 respectively. (See Figure 6.15.)
  A_k, for all k, is a sample from a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_A)$,
  B_k, for all k, is a sample from a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_B)$,
  C_k, for all k, is a sample from a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_C)$ and
  D_k, for all k, is a sample from a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_D)$.

- A_k, B_k, C_k and D_k are mutually independent for every $k \in \mathbb{N}$.

- (A_k, B_k, C_k, D_k) and (A_m, B_m, C_m, D_m) are independent for $k \neq m$.

- $x_1(k)$ for $k \in \mathbb{N}$ represents the time of the k-th firing of transition t_1,
  $x_2(k)$ for $k \in \mathbb{N}$ represents the time of the k-th firing of transition t_2.

- The initial conditions are such that the initial token distribution is as shown in Figure 6.15 and tokens have just started their holding time. $x_1(0) = -\infty$ and $x_2(0) = 0$.

The mathematical description of the evolution of the Wait & Switch system for $k \in \mathbb{N}$ is as follows:

$$\begin{aligned} x_1(k+1) &= \max\{\, x_1(k) \;+\; A_k,\; x_2(k) + B_k \,\}, \\ x_2(k+1) &= \max\{\, x_1(k+1) + C_k,\; x_2(k) + D_k \,\}. \end{aligned} \tag{6.31}$$

If we now define for $k \in \mathbb{N}^+$ the variables $b(k)$ and $f(k)$ as follows:

$$\begin{aligned} b(k) &\stackrel{\text{def}}{=} x_2(k) - x_1(k), \\ f(k) &\stackrel{\text{def}}{=} x_1(k+1) - x_2(k), \end{aligned} \tag{6.32}$$

then we can transform the evolution in (6.31) in the following more interesting description:

$$\begin{aligned} x_1(k+1) &= x_1(k) + \max\{\ A_k,\ b(k) + B_k\ \}, \\ x_2(k+1) &= x_2(k) + \max\{\ D_k,\ f(k) + C_k\ \}, \\ f(k) &= \max\{\ A_k - b(k),\ B_k\ \}, \\ b(k+1) &= \max\{\ D_k - f(k),\ C_k\ \}, \end{aligned} \tag{6.33}$$

or, equivalently, *totally de-coupled* as follows:

$$\begin{aligned} x_1(k+1) &= x_1(k) + \max\{\ A_k,\ b(k) + B_k\ \}, \\ b(k+1) &= \max\{\ D_k - \max[\ A_k - b(k),\ B_k\ ],\ C_k\ \}. \end{aligned} \tag{6.34}$$

in combination with

$$\begin{aligned} x_2(k+1) &= x_2(k) + \max\{\ D_k,\ f(k) + C_k\ \}, \\ f(k+1) &= \max\{\ A_{k+1} - \max[\ D_k - f(k),\ C_k\ ],\ B_{k+1}\ \}. \end{aligned} \tag{6.35}$$

The new descriptions in (6.34) and (6.35) are interesting because they *de-couple* the original system in two very simple subsystems. A figure of this de-coupling is shown in Figure 6.16. One could argue that this only is a "virtual" de-coupling because of the coupling between the variables $b(k)$ and $f(k)$. Note that the stochastic variables $b(k)$, A_k and B_k in one subsystem are mutually independent for a fixed k, and that the stochastic variables $f(k)$, C_k and D_k in the other subsystem are mutually independent for a fixed k as well. This means that if we "know" the stochastic processes of $b(k)$ and $f(k)$ for all $k \in \mathbb{N}$ then, although $b(k)$ and $f(k)$ are dependent, we can view each of the two smaller subsystems in Figure 6.16 as autonomous processes on themselves.

In order to obtain the distribution functions for $b(k)$ and $f(k)$ for $k \in \mathbb{N}$ we examine the last two equations from Expression (6.33) in more detail. If we define the distribution functions $F_{b(k)}(u)$ and $F_{f(k)}(u)$ for $u \in \mathbb{R}$ and the distribution functions for A_k, B_k, C_k

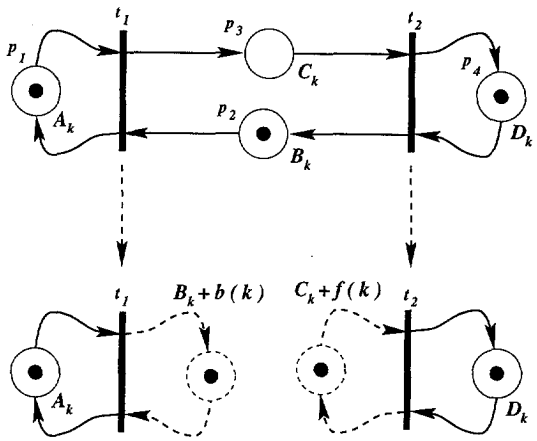

Figure 6.16: De-coupling of the Wait & Switch system.

and D_k as follows:

$$\begin{array}{rcl} F_{b(k)}(u) & \stackrel{\text{def}}{=} & \mathbb{P}\mathrm{r}(\ b(k) \leq u\ ), \\ F_{f(k)}(u) & \stackrel{\text{def}}{=} & \mathbb{P}\mathrm{r}(\ f(k) \leq u\ ), \\ F_A(u) & \stackrel{\text{def}}{=} & I\!P_A(\ A_k \leq u\ ), \\ F_B(u) & \stackrel{\text{def}}{=} & I\!P_B(\ B_k \leq u\ ), \\ F_C(u) & \stackrel{\text{def}}{=} & I\!P_C(\ C_k \leq u\ ), \\ F_D(u) & \stackrel{\text{def}}{=} & I\!P_D(\ D_k \leq u\ ), \end{array} \tag{6.36}$$

then, by means of the description in the last two equations of Expression (6.33), we can describe the evolution for $k \in I\!N$ of the distributions of $b(k)$ and $f(k)$ as follows:

$$\begin{array}{rcl} F_{f(k)}(u) & = & F_B(u) \int_{I\!R} F_A(s+u)\ \mathrm{d}F_{b(k)}(s), \\ F_{b(k+1)}(u) & = & F_C(u) \int_{I\!R} F_D(s+u)\ \mathrm{d}F_{f(k)}(s). \end{array} \tag{6.37}$$

In fact this completes the de-coupling for the Wait & Switch system because from Expression (6.37) we can derive everything we need of the stochastic processes for $b(k)$ and $f(k)$ for all $k \in I\!N$. By means of the the two sequences of distribution functions $F_{f(k)}(u)$ and $F_{b(k)}(u)$ we can derive *independently* the distribution functions for $x_1(k+1) - x_1(k)$ and $x_2(k+1) - x_2(k)$, which we will denote by $F_{t_1(k)}$ and $F_{t_2(k)}$ and define for $k \in I\!N$ as follows:

$$\begin{array}{rcl} F_{t_1(k)}(u) & \stackrel{\text{def}}{=} & \mathbb{P}\mathrm{r}(\ x_1(k+1) - x_1(k) \leq u\ ), \\ F_{t_2(k)}(u) & \stackrel{\text{def}}{=} & \mathbb{P}\mathrm{r}(\ x_2(k+1) - x_2(k) \leq u\ ). \end{array} \tag{6.38}$$

We derive, by means of the first two equations in Expression (6.33), that

$$\begin{aligned} F_{t_1(k)}(u) &= F_A(u) \int_{\mathbb{R}} F_B(u-s)\ \mathrm{d}F_{b(k)}(s), \\ F_{t_2(k)}(u) &= F_D(u) \int_{\mathbb{R}} F_D(u-s)\ \mathrm{d}F_{f(k)}(s), \end{aligned} \tag{6.39}$$

However, we are also interested in what happens if $k \to \infty$. We are interested in the limit (in case it exists) of the distribution functions $F_{f(k)}(u)$ and $F_{b(k)}(u)$ for $k \to \infty$. Therefore we will define:

$$\begin{aligned} F_b(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} F_{b(k)}(u), \\ F_f(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} F_{f(k)}(u). \end{aligned} \tag{6.40}$$

This means that these two limit distribution functions $F_b(u)$ and $F_f(u)$ are solutions to a fixed point problem under the transformations in Expression (6.37). The fixed point problem is to find two distribution functions $X(s)$ and $Y(s)$ such that the following two equations hold simultaneously:

$$\begin{aligned} X(u) &= F_B(u) \int_{\mathbb{R}} F_A(s+u)\ \mathrm{d}Y(s), \\ Y(u) &= F_C(u) \int_{\mathbb{R}} F_D(s+u)\ \mathrm{d}X(s). \end{aligned} \tag{6.41}$$

If we have found the distribution functions $F_f(u)$ and $F_b(u)$ we can easily find the stationary distribution of the time between two firings of transitions t_1 and t_2, $F_{t_1}(u)$ and $F_{t_2}(u)$ respectively, defined by:

$$\begin{aligned} F_{t_1}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} F_{t_1(k)}(u), \\ F_{t_2}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} F_{t_2(k)}(u). \end{aligned} \tag{6.42}$$

We derive by taking the limit in Expression (6.39) that

$$\begin{aligned} F_{t_1}(u) &= F_A(u) \int_{\mathbb{R}} F_B(u-s)\ \mathrm{d}F_b(s), \\ F_{t_2}(u) &= F_D(u) \int_{\mathbb{R}} F_D(u-s)\ \mathrm{d}F_f(s), \end{aligned} \tag{6.43}$$

which gives the desired distributions for the time between two firings of transitions t_1 and t_2.

Remark Note that the distribution functions of the time between two firings for transition t_1 and t_2 are not equal in general:

$$F_{t_1}(u) \neq F_{t_2}(u). \tag{6.44}$$

But the expectation of the time between two firings must be equal under all circumstances for both transitions:

$$\lim_{k\to\infty} \mathbb{E}\left[x_1(k+1) - x_1(k)\right] = \lim_{k\to\infty} \mathbb{E}\left[x_2(k+1) - x_2(k)\right], \tag{6.45}$$

or equivalently

$$\int_{\mathbb{R}} u \, \mathrm{d}F_{t_1}(u) = \int_{\mathbb{R}} u \, \mathrm{d}F_{t_2}(u). \tag{6.46}$$

This property is equivalent to the *conservation of flow* property mentioned in earlier chapters.

Application of the Wait & Switch system in larger chains

Suppose that we have a stochastic serial production line that can be described by means of one of the descriptions in (6.28), (6.29) or (6.30). Such a stochastic timed event graph is again shown in Figure 6.17. In this figure we have put the stochastic variables $T_{i,j}(k)$ that represent the stochastic holding time next to the places that correspond to it. (In Figure 6.17 we have omitted the counter k everywhere in order to make it more surveyable.) Suppose that for each position in the line we construct a separate Wait & Switch system that corresponds to it in a certain way also shown in Figure 6.17. We suppose that the behaviour of each position can be described by a Wait & Switch system. In order to construct the Wait & Switch model of the i-th position in the line we isolate the i-th position and add two "self loops" to it. In this way each of the two transitions t_i and t_{i+1} from position i receives one extra "self loop" represented by two places $p^f_{i,i}$ and $p^b_{i+1,i+1}$ with corresponding stochastic holding times $T^f_{i,i}(k)$ and $T^b_{i+1,i+1}(k)$ respectively. (Of course, again, the superscripts f and b stand for the first letters of the words *forward* and *backward* respectively.) The stochastic holding times $T^f_{i,i}(k)$ and $T^b_{i+1,i+1}(k)$ of each of these "self loops" are considered to represent the stochastic holding time of the "rest of the line" in a similar way as $B_k + b(k)$ and $C_k + f(k)$ represented the stochastic holding times of the "rest of the line" for the Wait & Switch model from Figure 6.16 that we have examined before.

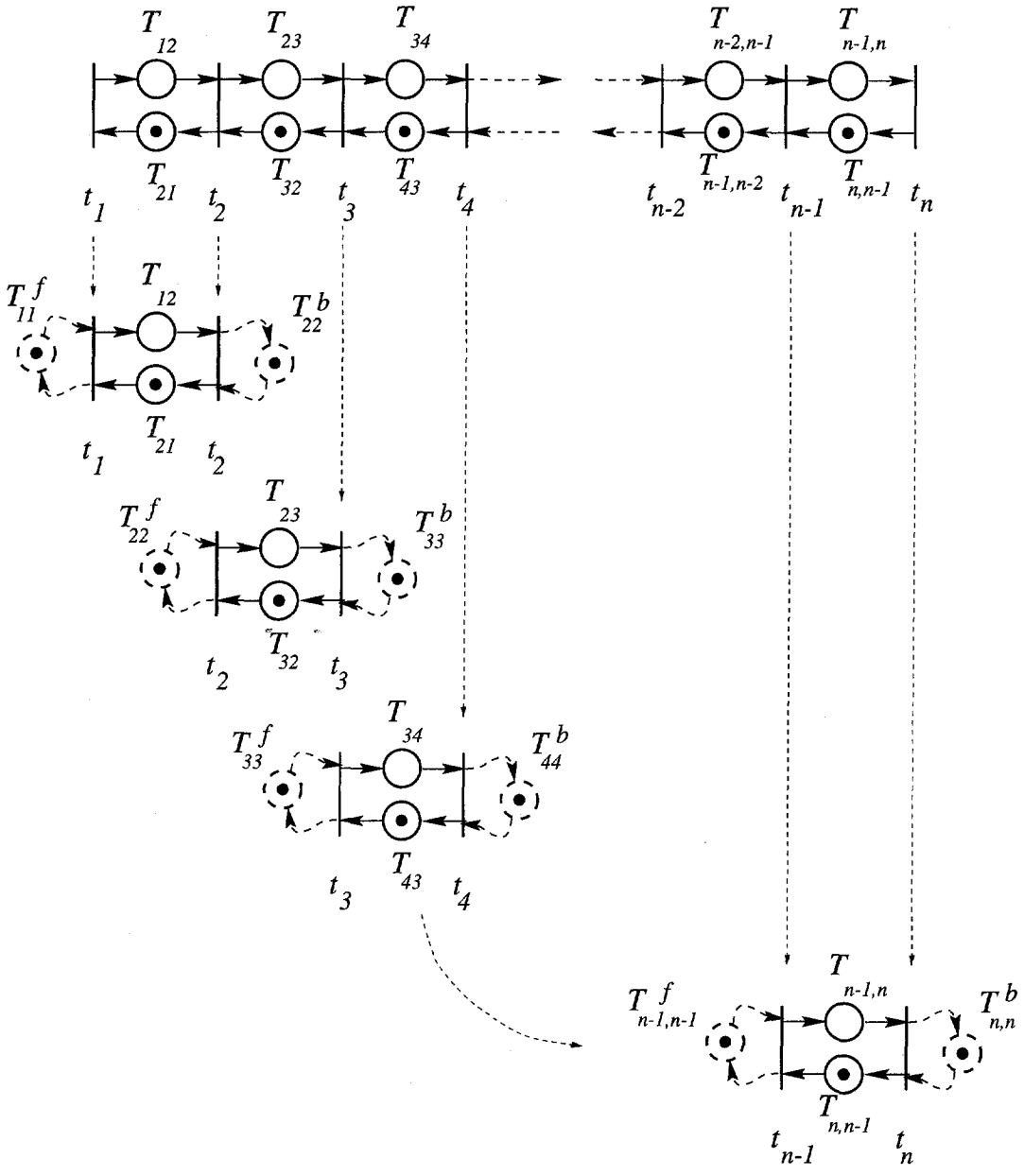

Figure 6.17: The timed event graph of a serial production line with blocking *decomposed* in several Wait & Switch systems.

If we define for $k \in I\!N$ and for $i, j \in \{1, 2, 3, ..., n\}$ that

$$\begin{aligned}
F_{T_{i,j}}(u) &\stackrel{\text{def}}{=} I\!P_{T_{i,j}}(\, T_{i,j}(k) \leq u \,), \quad \text{for } |i-j| = 1, \\
F_{T^b_{i,i}}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} I\!\Pr(\, T^b_{i,i}(k) \leq u \,), \\
F_{T^f_{i,i}}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} I\!\Pr(\, T^f_{i,i}(k) \leq u \,), \\
b_i(k) &\stackrel{\text{def}}{=} x_{i+1}(k) - x_i(k), \\
f_i(k) &\stackrel{\text{def}}{=} x_i(k+1) - x_{i+1}(k), \\
F_{b_i}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} I\!\Pr(\, b_i(k) \leq u \,), \\
F_{f_i}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} I\!\Pr(\, f_i(k) \leq u \,), \\
F_{t_i}(u) &\stackrel{\text{def}}{=} \lim_{k\to\infty} I\!\Pr(\, x_i(k+1) - x_i(k) \leq u \,).
\end{aligned} \tag{6.47}$$

then we can write down the equations for the i-th Wait & Switch model corresponding to the i-th position.

First we write down $F_{b_i}(u)$ and $F_{f_i}(u)$ as the solution the fixed point problem equivalent to Expression (6.39):

$$\begin{aligned}
F_{b_i}(u) &= F_{T_{i+1,i}}(u) \int_{I\!R} F_{T^f_{i,i}}(u+s)\, \mathrm{d}F_{f_i}(s), \\
F_{f_i}(u) &= F_{T_{i,i+1}}(u) \int_{I\!R} F_{T^f_{i+1,i+1}}(u+s)\, \mathrm{d}F_{b_i}(s).
\end{aligned} \tag{6.48}$$

Then we can formulate an expression for $F_{t_i}(u)$ and $F_{t_{i+1}}(u)$, similar to Expression (6.43):

$$\begin{aligned}
F_{t_i}(u) &= F_{T_{i,i}}(u) \int_{I\!R} F_{T_{i+1,i}}(u-s)\, \mathrm{d}F_{b_i}(s), \\
F_{t_{i+1}}(u) &= F_{T_{i+1,i+1}}(u) \int_{I\!R} F_{T_{i,i+1}}(u-s)\, \mathrm{d}F_{f_i}(s),
\end{aligned} \tag{6.49}$$

In order to "link" two neighbouring Wait & Switch models with numbers $i-1$ and i we define the following restrictions for linking:

$$\begin{aligned}
T^f_{i,i}(k) &\stackrel{\text{def}}{=} f_{i-1}(k) + T_{i-1,i}(k) \\
T^b_{i,i}(k) &\stackrel{\text{def}}{=} b_i(k) + T_{i+1,i}(k).
\end{aligned} \tag{6.50}$$

We assume that $f_{i-1}(k)$ and $T_{i-1,i}(k)$ are independent variables. We will also assume that $b_i(k)$ and $T_{i+1,i}(k)$ are independent variables. As a consequence from (6.50) we can derive that:

$$\begin{aligned} F_{T^f_{i,i}}(u) &= \int_{I\!R} F_{T_{i-1,i}}(u-s)\ \mathrm{d}F_{f_{i-1}}(s), \\ F_{T^b_{i,i}}(u) &= \int_{I\!R} F_{T_{i+1,i}}(u-s)\ \mathrm{d}F_{b_i}(s), \end{aligned} \tag{6.51}$$

The following conditions for $T^f_{1,1}$ and $T^b_{n,n}$ are considered as obvious *boundary conditions*:

$$T^f_{1,1}(k) = T^b_{n,n}(k) = 0, \quad \text{for all } k \in I\!N. \tag{6.52}$$

If we translate this in corresponding distribution functions we get:

$$F_{T^f_{1,1}}(u) = \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0, \end{cases} \quad , \qquad F_{T^b_{n,n}}(u) = \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0. \end{cases} \quad . \tag{6.53}$$

The equations in Expressions (6.48), (6.49) and (6.51), with the boundary conditions in (6.53), are the basic equations that we will try to solve in order to compute the distribution functions for the time between two firings of each transition t_i, $F_{t_i}(u)$ for $i \in \{1, 2, 3, ..., n\}$. If we have found a solution to the equations then, as a consequence of a property of all Wait & Switch systems described in Expression (6.46), we conclude that *conservation of flow* holds as follows for all $i, j \in \{1, 2, 3, ..., n\}$,

$$T \stackrel{\text{def}}{=} \int_{I\!R} u\ \mathrm{d}F_{t_i}(u) = \int_{I\!R} u\ \mathrm{d}F_{t_j}(u), \tag{6.54}$$

which is equivalent to the approximated average time needed by the production line to produce one single product. The average production rate denoted by R is then determined by:

$$R = \frac{1}{T}. \tag{6.55}$$

An algorithm to solve the equations.

We have found equations in Expressions (6.48), (6.49) and (6.51), with the boundary conditions in (6.53), but we did not yet describe a method to find a solution. In order to find a solution we suggest the following algorithm:

1. Take as initial conditions the following distribution functions for all $i \in \{1, 2, 3, ..., n-1\}$:

$$F^0_{T^f_{i,i}}(u) = \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0, \end{cases} \tag{6.56}$$

and the following distribution functions for all $i \in \{2, 3, ..., n\}$:

$$F^0_{T^b_{i,i}}(u) = \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0. \end{cases} \tag{6.57}$$

Initialize also the boundary conditions for $T^f_{1,1}$ and $T^b_{n,n}$ such that they hold for all iteration steps $s \in I\!N$:

$$F^s_{T^f_{1,1}}(u) = \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0, \end{cases} \qquad F^s_{T^b_{n,n}}(u) = \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0. \end{cases} \tag{6.58}$$

2. Initialize iteration step number s with $s := 1$.

3. **Forward aggregation**, do for $i = 1, 2, 3, ..., n-1$:

   a. Compute $F^{f,s}_{b_i}(u)$ and $F^{f,s}_{f_i}(u)$, the fixed point solution with counter i of the following equations that correspond to the equations in Expression (6.48):

$$\begin{aligned} F^{f,s}_{b_i}(u) &= F_{T_{i+1,i}}(u) \int_{I\!R} F^s_{T^f_{i,i}}(u+\tau)\, \mathrm{d}F^{f,s}_{f_i}(\tau), \\ F^{f,s}_{f_i}(u) &= F_{T_{i,i+1}}(u) \int_{I\!R} F^{s-1}_{T^f_{i+1,i+1}}(u+\tau)\, \mathrm{d}F^{f,s}_{b_i}(\tau). \end{aligned} \tag{6.59}$$

   b. Compute $F^s_{T^f_{i+1,i+1}}(u)$ by means of the first equation in Expression (6.51) and the solution $F^s_{f_i}(u)$ found in a. as follows:

$$F^s_{T^f_{i+1,i+1}}(u) = \int_{I\!R} F_{T_{i,i+1}}(u-\tau)\, \mathrm{d}F^s_{f_i}(\tau) \tag{6.60}$$

4. **Backward aggregation**, do for $i = n-1, n-2, n-3, ..., 1$:

   a. Compute $F^{b,s}_{b_i}(u)$ and $F^{b,s}_{f_i}(u)$, the fixed point solution with counter i of the following equations that correspond to the equations in Expression (6.48):

$$\begin{aligned} F^{b,s}_{b_i}(u) &= F_{T_{i+1,i}}(u) \int_{I\!R} F^s_{T^f_{i,i}}(u+\tau)\, \mathrm{d}F^{b,s}_{f_i}(\tau), \\ F^{b,s}_{f_i}(u) &= F_{T_{i,i+1}}(u) \int_{I\!R} F^s_{T^f_{i+1,i+1}}(u+\tau)\, \mathrm{d}F^{b,s}_{b_i}(\tau). \end{aligned} \tag{6.61}$$

   b. Compute $F^s_{T^b_{i,i}}(u)$ by means of the second equation in Expression (6.51) and the solution $F^s_{b_i}(u)$ found in a. as follows:

$$F^s_{T^b_{i,i}}(u) = \int_{I\!R} F_{T_{i+1,i}}(u-\tau)\, \mathrm{d}F^s_{b_i}(\tau) \tag{6.62}$$

5. If sufficient convergence is *not* reached then increase iteration counter s with $s := s+1$ and go to 3. If sufficient convergence is reached then continue here and go to 6.

6. Compute $F^s_{t_i}(u)$ for $i = 1, 2, 3, ..., n$ by means of the following equations that correspond to equations from Expression (6.49) as follows:

$$F^s_{t_i}(u) \;=\; F^s_{T^f_{i,i}}(u) \int_{\mathbb{R}} F_{T_{i+1,i}}(u-s)\, \mathrm{d}F^{b,s}_{b_i}(s). \tag{6.63}$$

7. Stop.

Remark We did not describe a method to find the fixed point distributions as described in 3.a and 4.a. This fixed point problem is similar to the fixed point problem as described in Expression (6.41). A straightforward way to find the fixed point solution in (6.41) is to start with the initial condition

$$F_{b(0)}(u) \;=\; \begin{cases} 0, & \text{for } u < 0, \\ 1, & \text{for } u \geq 0. \end{cases}, \tag{6.64}$$

and continue the iteration for $k = 1, 2, 3, ...$ in Expression (6.37) until sufficient convergence is reached.

Note also that it is not necessary to find an exact solution each time we perform steps 3.a and 4.a and that we can re-use the solutions of steps 3.a and 4.a from iteration cycle s for the same steps in the next iteration cycle $s + 1$. If we play with these ideas we can come up with algorithms that are much more efficient. The purpose of this algorithm, however, is just to offer a method to compute the solution to our equations. Efficiency was not our aim.

6.3 Conclusions

In this chapter we introduced the concept of the stochastic timed event graph. First we described the basic properties of such graphs. Then, for the purpose of illustration, we treated a simple two-dimensional autonomous stochastic timed event graph. In spite of its simplicity, however, we were not able to solve the stationary behaviour of this simple two-dimensional system in the most general form. This illustrates the complexity of these graphs and shows that explicit expressions for the stationary distribution of such graphs is still beyond our reach.

After this illustration of the complexity we give a very convenient description of serial queueing networks by means of stochastic timed event graphs. This *most elegant* description of serial queueing systems is *more general* than the descriptions that come from standard queueing theory because it allows each position in a queue to have arbitrary

transportation time distributions, it allows every possible collection of waiting time distributions for the servers and it allows also an arbitrary distributed repair time between two services. This new approach does *not* make a fundamental difference between parts waiting inside a server and parts waiting at a position in a queue: positions in queues are dealt with in exactly the same way as positions inside servers. In fact the new description considers each single position as being a small server in itself. Some examples of how production lines from previous chapters can be described by means of event graphs were given.

This elegant description by means of stochastic timed event graphs is useless if it does not lead to a method my means of which we can solve its stationary behaviour. It appears possible to *decompose* the serial queueing system into smaller graphs that we denote by the term *Wait & Switch System.* By means of thorough examination of the small Wait & Switch Systems we can solve their stationary distributions. The decomposition of the serial queueing network in Wait & Switch Systems, of which we can solve the stationary distributions, allows us to find expressions for the stationary distribution of the total network. Then an alternate forward and backward iteration algorithm is described by means of which we can find an estimate that satisfies all the restrictions in the expressions. The solution from the algorithm allows us to estimate the overall average production rate and any possible waiting time distribution that we are interested in.

Results from the algorithm are not given in this chapter. This is because it is our aim to find an *exact* solution method the stationary distributions. It can be shown that the solution to the restrictions that we find by means of the algorithm is not an exact solution to the stationary distribution of the overall queueing network. Therefore it is more interesting to look for better methods than to examine the solution from the algorithm.

The descriptions and the methods that we describe here offer a brand new approach to serial queueing networks with blocking. It is a promising method that certainly needs further research in a near future.

Appendix A

Used Functions and their Properties

In Chapter 2 and Appendix B we use the rather complicated functions Ψ_N, R_N, $\Omega_{c,N}$ and Q_N. In this appendix we will describe these functions and their properties in full detail. Each of the functions will be described in a separate section.

A.1 The Function Ψ_N and Its Properties

The function Ψ_N is introduced in Expression (2.15) and is used throughout Chapter 2 and Appendix B as one of the basic functions.

Definition 1

$$\Psi_N \ : \ [0,1] \times [0,1] \ \mapsto \ [0,1],$$

$$\Psi_N(x,y) \stackrel{\text{def}}{=} \begin{cases} \dfrac{1-\alpha}{1-\dfrac{y}{x}\alpha^{N+1}}, & \text{for} \quad (x,y) \in (0,1)^2 \text{ with } x \neq y \ , \\ \dfrac{1}{N+x}, & \text{for} \quad (x,y) \in [0,1]^2 \text{ with } x = y \ , \\ \max(2-N,0).(1-x), & \text{for} \quad (x,y) \in [0,1]^2 \text{ with } \dfrac{y(1-x)}{y-x} = 0 \ , \\ 1, & \text{for} \quad (x,y) \in [0,1]^2 \text{ with } \dfrac{x(1-y)}{x-y} = 0 \ , \end{cases} \tag{A.1}$$

where

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

Theorem 2 (Ψ_N property 1)

$$\Psi_N(x,y) = \frac{1}{y} \cdot \frac{x}{x+\alpha+\alpha^2+...+\alpha^N}, \quad \forall (x,y) \in (0,1)^2, \tag{A.2}$$

where

$$\alpha(x,y) = \frac{x\,(1-y)}{(1-x)\,y}.$$

Proof. First we observe that:

$$x = y \quad \Leftrightarrow \quad \alpha(x,y) = 1\ .$$

In case $(x,y) \in (0,1)^2$ and $x = y$ then the right side of (A.2) reduces to

$$\frac{1}{x} \cdot \frac{x}{x+1+1^2+...+1^N} = \frac{1}{N+x} = \Psi_N(x,x),$$

which is in accordance with definition (A.1).

In case $(x,y) \in (0,1)^2$ and $x \neq y$ then, since $\alpha \neq 1$, the right side of (A.2) can be rewritten as follows:

$$\begin{aligned} \frac{1}{y} \cdot \frac{x}{x+\alpha+\alpha^2+...+\alpha^N} &= \frac{x}{y} \cdot \frac{1-\alpha}{1-\alpha} \cdot \frac{1}{x+\alpha+\alpha^2+...+\alpha^N} \\ &= \frac{x}{y} \cdot \frac{1-\alpha}{(1-\alpha).x+\alpha-\alpha^{N+1}}. \end{aligned}$$

Since $(1-\alpha).x + \alpha = \frac{x}{y}$ we conclude that

$$\begin{aligned} \frac{x}{y} \cdot \frac{1-\alpha}{(1-\alpha).x+\alpha-\alpha^{N+1}} &= \frac{x}{y} \cdot \frac{1-\alpha}{\frac{x}{y}-\alpha^{N+1}} \\ &= \frac{1-\alpha}{1-\frac{y}{x}\alpha^{N+1}} \\ &= \Psi_N(x,y), \end{aligned}$$

which is also in accordance with definition (A.1).
This completes the proof of "Ψ_N property 1".

■

Theorem 3 (Ψ_N Property 2) *The functions Ψ_N satisfy the following recurrent relation in $N \in \mathbb{N}$ for all $(x,y) \in (0,1)^2$:*

$$\begin{aligned} \Psi_0(x,y) &= \frac{1}{y}, \\ \Psi_{N+1}(x,y) &= \frac{\Psi_N(x,y)}{\Psi_N(x,y)+\alpha}, \end{aligned} \tag{A.3}$$

where

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

Proof. First of all it is easy to check from the definition that $\Psi_0(x,y) = \frac{1}{y}$.
Now let $(x,y) \in (0,1)^2$ and

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

Then by using "Ψ_N Property 1" we conclude that

$$\begin{aligned} \frac{\Psi_N(x,y)}{\Psi_N(x,y)+\alpha} &= \frac{1}{1+\frac{\alpha}{\Psi_N(x,y)}} \\ &= \frac{1}{1+\alpha\cdot\frac{y}{x}\cdot(x+\alpha+...+\alpha^N)} \\ &= \frac{1}{y}\cdot\frac{x}{\frac{x}{y}+\alpha x+\alpha(\alpha^2+...+\alpha^N)}. \end{aligned}$$

Since $\frac{x}{y} = \alpha + (1-\alpha).x$ we conclude that

$$\begin{aligned} \frac{1}{y}\cdot\frac{x}{\frac{x}{y}+\alpha x+\alpha(\alpha^2+...+\alpha^N)} &= \frac{1}{y}\cdot\frac{x}{\alpha+(1-\alpha)x+\alpha x+\alpha(\alpha^2+...+\alpha^N)} \\ &= \frac{1}{y}\cdot\frac{x}{x+\alpha+\alpha^2+...+\alpha^{N+1}} \\ &= \Psi_{N+1}(x,y). \end{aligned}$$

The last step uses again "Ψ_N Property 1".
This completes the proof of "Ψ_N Property 2".

■

Theorem 4 (Ψ_N Property 3) *Ψ_N is continuous on $[0,1]^2$ except for points $(0,0)$ and $(1,1)$.*

Proof. It is easy to see that $\Psi_N(x,0) = 0$ and $\Psi_N(0,x) = 1$ for all $x \in (0,1]$. This implies that Ψ_N is discontinuous in $(0,0)$.
Similarly we notice that $\Psi_N(x,1) = 1$ and $\Psi_N(1,x) = 0$ for all $x \in [0,1)$ which implies discontinuity in $(1,1)$.

■

Theorem 5 (Ψ_N Property 4) *Ψ_N and all its partial derivatives of any order are continuous on $(0,1)^2$.*

A.2 The Function R_N and Its Properties

The function R_N is introduced in Expression (2.20) and is used throughout Chapter 2 and Appendix B as one of the basic functions.

Definition 2

$$R_N : [0,1] \times [0,1] \mapsto [0,1],$$
$$R_N(x,y) \stackrel{\text{def}}{=} x \cdot (1 - (1-y)\Psi_N(1-x, 1-y)). \tag{A.4}$$

Theorem 6 (R_N property 1)

$$R_N(x,y) = x \cdot \left\{ 1 - \frac{1-x}{1-x+\beta+\beta^2+...+\beta^N} \right\}, \quad \forall (x,y) \in (0,1)^2, \tag{A.5}$$

where

$$\beta(x,y) = \frac{y\,(1-x)}{(1-y)\,x}.$$

Proof. This follows directly from the definition of R_N and "Ψ_N property 1".

■

Theorem 7 (R_N property 2)

$$R_N(x,y) = \begin{cases} x \cdot \dfrac{\beta^{N+1} - \beta}{\beta^{N+1} - \dfrac{1-x}{1-y}}, & \forall (x,y) \in (0,1)^2 \text{ with } x \neq y \\ \dfrac{N\,x}{N+1-x}, & \forall (x,y) \in (0,1)^2 \text{ with } x = y \\ xy, & \forall (x,y) \in [0,1]^2 \text{ with } xy(1-x)(1-y) = 0 \end{cases} . \tag{A.6}$$

where

$$\beta(x,y) = \frac{y\ (1-x)}{(1-y)\ x}.$$

Proof. Let $(x,y) \in (0,1)^2$ and let

$$\beta(x,y) = \frac{y\ (1-x)}{(1-y)\ x}.$$

Then follows from "R_N property 1" that

$$R_N(x,y) \ = \ x \cdot \frac{\beta + \beta^2 + ... + \beta^N}{1 - x + \beta + \beta^2 + ... + \beta^N}.$$

In case that $x = y$ then $\beta = 1$ and $R_N(x,y)$ reduces to

$$\begin{aligned} R_N(x,y) \ &= \ x \cdot \frac{1 + 1^2 + ... + 1^N}{1 - x + 1 + 1^2 + ... + 1^N} \\ &= \ \frac{N\ x}{N + 1 - x}, \end{aligned}$$

which is in accordance with "R_N property 2".

In case that $x \neq y$ then $\beta \neq 1$ and we can conclude that

$$\begin{aligned} x \cdot \frac{\beta + \beta^2 + ... + \beta^N}{1 - x + \beta + \beta^2 + ... + \beta^N} \ &= \ x \cdot \frac{\beta - 1}{\beta - 1} \cdot \frac{\beta + \beta^2 + ... + \beta^N}{1 - x + \beta + \beta^2 + ... + \beta^N} \\ &= \ x \cdot \frac{\beta^{N+1} - \beta}{\beta^{N+1} - \{\beta + (1-x)(1-\beta)\}}. \end{aligned}$$

If we now use that $\beta + (1-\beta)(1-x) = \frac{1-x}{1-y}$ we get that

$$x \cdot \frac{\beta^{N+1} - \beta}{\beta^{N+1} - \{\beta + (1-x)(1-\beta)\}} \ = \ x \cdot \frac{\beta^{N+1} - \beta}{\beta^{N+1} - \frac{1-x}{1-y}}$$

This ends the proof of "R_N property 2" for the interior $(x,y) \in (0,1)^2$.
For the edge expressed by all $(x,y) \in [0,1]^2$ with $xy(1-x)1-y) = 0$ it is simple to check the result. This completes the proof of "R_N property 2".

■

Theorem 8 (R_N **property 3)** *The functions R_N satisfy the following recurrent relation in $N \in \mathbb{N}$ for all $(x,y) \in (0,1)^2$:*

$$\begin{aligned} R_0(x,y) &= 0, \\ R_{N+1}(x,y) &= x \cdot y \cdot \frac{1 - R_N(x,y)}{x + y - xy - R_N(x,y)}. \end{aligned} \tag{A.7}$$

Proof. Let $(x,y) \in (0,1)^2$. First of all it is easy to check from the definition that $R_0(x,y) = 0$.
From the definition of R_N follows that

$$R_{N+1}(x,y) = x(1 - (1-y)\Psi_{N+1}(1-x, 1-y)).$$

With "Ψ_N property 2" this transforms to

$$R_{N+1}(x,y) = x(1 - (1-y)\frac{\Psi_N(1-x,1-y)}{\Psi_N(1-x,1-y) + \beta}), \tag{A.8}$$

where $\beta(x,y) = \frac{y\,(1-x)}{x\,(1-y)}$.
If we transform the definition of R_N such that we get an explicit formula for $\Psi_N(1-x,1-y)$ in terms of $R_N(x,y)$ we obtain

$$\Psi_N(1-x,1-y) = \frac{x - R_N(x,y)}{x(1-y)}.$$

Using the last expression for $\Psi_N(1-x,1-y)$ we get that

$$\begin{aligned} \frac{\Psi_N(1-x,1-y)}{\Psi_N(1-x,1-y) + \beta} &= \frac{\frac{x - R_N(x,y)}{x(1-y)}}{\frac{x - R_N(x,y)}{x(1-y)} + \frac{y\,(1-x)}{x\,(1-y)}} \\ &= \frac{x - R_N(x,y)}{x - R_N(x,y) + y(1-x)} \\ &= \frac{x - R_N(x,y)}{x + y - xy - R_N(x,y)}. \end{aligned}$$

Now we continue by substituting this in Expression (A.8) and we get that

$$\begin{aligned} R_{N+1}(x,y) &= x(1 - (1-y)\frac{x - R_N(x,y)}{x + y - xy - R_N(x,y)}) \\ &= x \cdot y \cdot \frac{1 - R_N(x,y)}{x + y - xy - R_N(x,y)}. \end{aligned}$$

This completes the proof of "R_N property 3". ∎

Theorem 9 (R_N **property 4)**

$$R_N(x,y) = R_N(y,x), \qquad \forall (x,y) \in [0,1]^2 \tag{A.9}$$

Proof. In case that $(x,y) \in (0,1)^2$ and $x \neq y$ let

$$\beta(x,y) = \frac{y\,(1-x)}{(1-y)\,x}.$$

Then follows from "R_N property 2" that for $x \neq y$

$$\begin{aligned}
R_N(x,y) &= x \cdot \frac{\beta^{N+1}(x,y) - \beta(x,y)}{\beta^{N+1}(x,y) - \dfrac{1-x}{1-y}} \\
&= x \cdot \frac{\beta(x,y) - \beta^{N+1}(x,y)}{\dfrac{1-x}{1-y} - \beta^{N+1}(x,y)} \cdot \frac{\dfrac{1-y}{1-x} \cdot \beta^{-(N+1)}(x,y)}{\dfrac{1-y}{1-x} \cdot \beta^{-(N+1)}(x,y)} \\
&= y \cdot \frac{\dfrac{x\,(1-y)}{(1-x)\,y}\ (\beta^{-N}(x,y) - 1)}{\beta^{-(N+1)}(x,y) - \dfrac{1-y}{1-x}} \\
&= y \cdot \frac{\beta^{-(N+1)}(x,y) - \beta^{-1}(x,y)}{\beta^{-(N+1)}(x,y) - \dfrac{1-y}{1-x}} \\
&= y \cdot \frac{\beta^{N+1}(y,x) - \beta(y,x)}{\beta^{N+1}(y,x) - \dfrac{1-y}{1-x}} \\
&= R_N(y,x).
\end{aligned}$$

The second last step in the proof uses the fact that $\beta(y,x) = \beta^{-1}(x,y)$, the last step uses again "R_N property 2".

The case $x = y$ obviously is not important for this symmetry property.
What is left to check are the cases in which x or y are either 0 or 1.
From "R_N property 2" follows that $R_N(x,y) = xy = R_N(y,x)$ in these cases.
This completes the proof of "R_N property 4".

■

Theorem 10 (R_N **property 5)**

$$0 < R_N(x,y) < x, \qquad \forall (x,y) \in (0,1)^2. \tag{A.10}$$

Proof. Let $(x, y) \in (0,1)^2$ and let

$$\beta(x,y) = \frac{y\,(1-x)}{(1-y)\,x}.$$

Then obviously $\beta > 0$ and $1 - x > 0$ and therefore also

$$\frac{1-x}{1-x+\beta+\beta^2+...+\beta^N} > 0.$$

And we conclude that

$$\begin{aligned}
-\frac{1-x}{1-x+\beta+\beta^2+...+\beta^N} &< 0, \\
1-\frac{1-x}{1-x+\beta+\beta^2+...+\beta^N} &< 1, \\
x\,.\left\{1-\frac{1-x}{1-x+\beta+\beta^2+...+\beta^N}\right\} &< x, \\
R_N(x,y) &< x.
\end{aligned}$$

The last step uses "R_N property 1".

On the other hand again since $\beta > 0$, $1 - x > 0$ and $x > 0$

$$\begin{aligned}
x\cdot\frac{\beta+\beta^2+...+\beta^N}{1-x+\beta+\beta^2+...+\beta^N} &> 0, \\
x\,.\left\{1-\frac{1-x}{1-x+\beta+\beta^2+...+\beta^N}\right\} &> 0, \\
R_N(x,y) &> 0.
\end{aligned}$$

The last step again uses "R_N property 1".
This completes the proof of "R_N property 5".

■

Theorem 11 **(R_N property 6)** *R_N and all its partial derivatives of any order exist and are continuous on $(0,1)^2$.*

Theorem 12 **(R_N property 7)**

$$\frac{\partial R_N}{\partial x}(x,y) = \left(\frac{1-y}{1-x}\right)^2 \cdot \frac{1+2\alpha+3\alpha^2+4\alpha^3+...+N\alpha^{N-1}}{\left(1-y+\alpha+\alpha^2+\alpha^3+...+\alpha^N\right)^2}, \quad \forall (x,y) \in (0,1)^2, \tag{A.11}$$

where

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

Proof. Let $(x,y) \in (0,1)^2$ and

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

A combination of "R_N property 1" and "R_N property 4" makes that

$$R_N(x,y) \quad = \quad y\ .\ \left\{\ 1\ -\ \frac{1-y}{1-y+\alpha+\alpha^2+...+\alpha^N}\ \right\}.$$

If we differentiate both sides to the variable x we obtain

$$\begin{aligned}
\frac{\partial R_N}{\partial x}(x,y) &= -y\ (1-y)\cdot\frac{\partial}{\partial x}\left(\frac{1}{1-y+\alpha+\alpha^2+...+\alpha^N}\right), \\
&= y\ (1-y)\cdot\frac{\dfrac{\partial}{\partial x}\left(1-y+\alpha+\alpha^2+...+\alpha^N\right)}{\left(1-y+\alpha+\alpha^2+\alpha^3+...+\alpha^N\right)^2}, \\
&= y\ (1-y)\cdot\frac{\dfrac{d}{d\alpha}\left(\alpha+\alpha^2+...+\alpha^N\right)\cdot\dfrac{\partial\alpha}{\partial x}}{\left(1-y+\alpha+\alpha^2+\alpha^3+...+\alpha^N\right)^2}, \\
&= y\ (1-y)\cdot\frac{\left(1+2\alpha+3\alpha^2+...+N\alpha^{N-1}\right)\cdot\dfrac{1-y}{y(1-x)^2}}{\left(1-y+\alpha+\alpha^2+\alpha^3+...+\alpha^N\right)^2}, \\
&= \left(\frac{1-y}{1-x}\right)^2\cdot\frac{\left(1+2\alpha+3\alpha^2+...+N\alpha^{N-1}\right)}{\left(1-y+\alpha+\alpha^2+\alpha^3+...+\alpha^N\right)^2}.
\end{aligned}$$

This completes the proof of "R_N property 7".

■

Theorem 13 (R_N **property 8**)

$$0 < \frac{\partial R_N}{\partial x}(x,y) < \frac{R_N(x,y)}{x}, \qquad \forall (x,y) \in (0,1)^2. \tag{A.12}$$

Proof. Let $(x,y) \in (0,1)^2$ and

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y},$$

then obviously $\alpha > 0$, $1 - y > 0$ and $1 - x > 0$ and therefore

$$\left(\frac{1-y}{1-x}\right)^2 \cdot \frac{\left(1 + 2\alpha + 3\alpha^2 + ... + N\alpha^{N-1}\right)}{\left(1 - y + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^N\right)^2} > 0,$$

$$\frac{\partial R_N}{\partial x}(x, y) > 0.$$

This is a direct consequence from "R_N property 7".
The other inequality in "R_N property 8" is more difficult. In order to show that $\frac{\partial R_N}{\partial x}(x, y) < \frac{R_N(x, y)}{x}$ we consider 3 cases for $(x, y) \in (0, 1)^2$: $x < y$, $x > y$ and $x = y$.

Let $(x, y) \in (0, 1)^2$ and $x < y$.
Then we use that $\alpha^i < 1$, $\forall i \in I\!N^+$, $1 - \alpha > 0$, and $\frac{1-x}{1-y} \cdot \alpha = \frac{x}{y} < 1$ and derive

$$1 + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^{N-1} \leq N,$$

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^{N-1}) < N,$$

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^{N-1})(1 - \alpha) < N(1 - \alpha),$$

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 - \alpha^N) < N(1 - \alpha).$$

We want to derive the same result for the case $(x, y) \in (0, 1)^2$ and $x < y$.
Therefore we use that $\alpha^i > 1$, $\forall i \in I\!N^+$, $1 - \alpha < 0$, and $\frac{1-x}{1-y} \cdot \alpha = \frac{x}{y} > 1$ and derive

$$1 + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^{N-1} \geq N,$$

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^{N-1}) > N,$$

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 + \alpha + \alpha^2 + \alpha^3 + ... + \alpha^{N-1})(1 - \alpha) < N(1 - \alpha),$$

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 - \alpha^N) < N(1 - \alpha).$$

We conclude that the inequality

$$\frac{1-x}{1-y} \cdot \alpha \cdot (1 - \alpha^N) < N(1 - \alpha),$$

holds for all $(x,y) \in (0,1)^2$ with $x \neq y$.

Now let us therefore combine the cases $x > y$ and $x < y$ and consider the case that $(x,y) \in (0,1)^2$ with $x \neq y$. Then we continue with what we have just derived

$$\begin{aligned}
\frac{1-x}{1-y} \cdot \alpha \cdot (1-\alpha^N) &< N(1-\alpha), \\
-\frac{1-x}{1-y} \cdot \alpha^{N+1} \cdot (1-\alpha^N) &> -N\alpha^N(1-\alpha), \\
(1-\alpha^N) - \frac{1-x}{1-y} \cdot \alpha^{N+1} \cdot (1-\alpha^N) &> (1-\alpha^N) - N\alpha^N(1-\alpha), \\
(1-\alpha^N)(1 - \frac{1-x}{1-y} \cdot \alpha^{N+1}) &> 1 - (N+1)\alpha^N + N\alpha^{N+1}, \\
(1-\alpha^N)(\frac{1-y}{1-x} - \alpha^{N+1}) &> \frac{1-y}{1-x} \cdot (1 - (N+1)\alpha^N + N\alpha^{N+1}),
\end{aligned}$$

For the following step we use the fact that $\frac{1-y}{1-x} = (1-\alpha)(1-y) + \alpha$ and so

$$\begin{aligned}
(1-\alpha^N)(\frac{1-y}{1-x} - \alpha^{N+1}) &> \frac{1-y}{1-x} \cdot (1 - (N+1)\alpha^N + N\alpha^{N+1}), \\
(1-\alpha^N)\left\{(1-\alpha)(1-y) + \alpha - \alpha^{N+1}\right\} &> \frac{1-y}{1-x} \cdot (1 - (N+1)\alpha^N + N\alpha^{N+1}), \\
\frac{1-\alpha^N}{1-\alpha} \cdot \frac{(1-\alpha)(1-y) + \alpha - \alpha^{N+1}}{1-\alpha} &> \frac{1-y}{1-x} \cdot \frac{1 - (N+1)\alpha^N + N\alpha^{N+1}}{(1-\alpha)^2}.
\end{aligned}$$

After some calculus you can conclude that this is exactly similar to

$$\begin{aligned}
&(1 + \alpha + \alpha^2 + ... + \alpha^{N-1}) \cdot ... \\
&\qquad (1 - y + \alpha + \alpha^2 + ... + \alpha^N) > \frac{1-y}{1-x}(1 + 2\alpha + 3\alpha^2 + ... + N\alpha^{N-1}), \\
&\frac{1-y}{1-x} \cdot \frac{1 + \alpha + \alpha^2 + ... + \alpha^{N-1}}{1 - y + \alpha + \alpha^2 + ... + \alpha^N} > \left(\frac{1-y}{1-x}\right)^2 \cdot \frac{1 + 2\alpha + 3\alpha^2 + ... + N\alpha^{N-1}}{(1 - y + \alpha + \alpha^2 + ... + \alpha^N)^2}.
\end{aligned}$$

On the left side we now use that $\frac{1-y}{1-x} = \frac{y}{x} \cdot \alpha$ and we get

$$\frac{y}{x} \cdot \frac{\alpha + \alpha^2 + \alpha^3 + ... + \alpha^N}{1 - y + \alpha + \alpha^2 + ... + \alpha^N} > \left(\frac{1-y}{1-x}\right)^2 \cdot \frac{1 + 2\alpha + 3\alpha^2 + ... + N\alpha^{N-1}}{(1 - y + \alpha + \alpha^2 + ... + \alpha^N)^2},$$

$$\frac{y}{x}\cdot\left\{1-\frac{1-y}{1-y+\alpha+\alpha^2+...+\alpha^N}\right\} > \left(\frac{1-y}{1-x}\right)^2\cdot\frac{1+2\alpha+3\alpha^2+...+N\alpha^{N-1}}{(1-y+\alpha+\alpha^2+...+\alpha^N)^2},$$

$$\frac{R_N(y,x)}{x} > \frac{\partial R_N}{\partial x}(x,y),$$

$$\frac{R_N(x,y)}{x} > \frac{\partial R_N}{\partial x}(x,y).$$

The last two steps use "R_N property 1", "R_N property 7" and "R_N property 4" respectively. Now we have shown that $\frac{R_N(x,y)}{x} > \frac{\partial R_N}{\partial x}(x,y)$ for all $(x,y)\in(0,1)^2$ and $x\neq y$. What is left to prove is that $\frac{R_N(x,y)}{x} > \frac{\partial R_N}{\partial x}(x,y)$ also holds for $(x,y)\in(0,1)^2$ and $x=y$. If we use "R_N property 2" for $x=y$ then we get

$$\frac{R_N(x,x)}{x} = \frac{\frac{N\,x}{N+1-x}}{x},$$

$$= \frac{N}{N+1-x}.$$

And if we use "R_N property 7" for $x=y$ we get

$$\frac{\partial R_N}{\partial x}(x,x) = \left(\frac{1-x}{1-x}\right)^2\cdot\frac{1+2.1+3.1^2+4.1^3+...+N.1^{N-1}}{(1-x+1+1^2+1^3+...+1^N)^2},$$

$$= \frac{1+2+3+4+...+N}{(1-x+1+1+1+...+1)^2},$$

$$= \frac{\frac{1}{2}N(N+1)}{(N+1-x)^2},$$

To check that in this case also $\frac{R_N(x,y)}{x} > \frac{\partial R_N}{\partial x}(x,y)$ we examine therefore

$$\frac{N}{N+1-x} > \frac{\frac{1}{2}N(N+1)}{(N+1-x)^2},$$

$$2N(N+1-x) > N(N+1),$$

$$\frac{1}{2}N(N+1) > x,$$

which obviously holds for all $x\in(0,1)$ and all $N\in\mathbb{N}^+$.
This completes the proof of "R_N property 8".

■

The author is grateful to his colleague Peter Sonneveld, a fellow researcher who works and teaches at the Delft University of Technology at the department of "Applied Mathematics and Computer Science". Mr Sonneveld has put a lot of effort in finding the proof of the following theorem denoted by "R_N property 9". It is a version of his proof that follows the theorem hereafter. Ideas from his proof have led to the derivation of the recursive properties described in "Ψ_N property 2" and "R_N property 3" as well.

Theorem 14 (R_N property 9) *For all $N \in \mathbb{N}^+$ and $(x,y) \in (0,1)^2$ the following holds:*

$$\begin{aligned} \frac{\partial^2 R_N}{\partial x^2}(x,y) &< 0, \\ \frac{\partial^2 R_N}{\partial y^2}(x,y) &< 0. \end{aligned} \tag{A.13}$$

Proof. Let a function $\psi_{y,N}$ for $y \in (0,1)$ be defined as follows:

$$\begin{aligned} \psi_{y,N} &: (0,\infty) \mapsto (0,1), \\ \psi_{y,N}(\alpha) &= \frac{1-y+y\alpha}{1-y+\alpha+\alpha^2+...+\alpha^N} \end{aligned} \tag{A.14}$$

Then, since $1-y+y\alpha = \frac{1-y}{1-x}$, it can easily be shown by means of "Ψ_N property 1" that

$$\psi_{y,N}\left(\frac{x\ (1-y)}{(1-x)\ y}\right) = \Psi_N(1-y,1-x).$$

Then, as a consequence of "Ψ_N property 2", we can conclude that

$$\psi_{y,N+1}(\alpha) = \frac{\psi_{y,N}(\alpha)}{\psi_{y,N}(\alpha)+\alpha}.$$

A combination of the definition of R_N and "R_N property 4" has the following result:

$$\begin{aligned} \frac{\partial^2 R_N}{\partial x^2}(x,y) &= \frac{\partial^2}{\partial x^2}\left(y\cdot(1-(1-x)\Psi_N(1-y,1-x))\right), \\ &= -y.\frac{\partial^2}{\partial x^2}\left((1-x)\psi_{y,N}(\alpha)\right), \end{aligned}$$

where

$$\alpha(x,y) = \frac{x\ (1-y)}{(1-x)\ y}.$$

If we continue we find a remarkable result, found by Mr Sonneveld:

$$\begin{aligned}
-y.\frac{\partial^2}{\partial x^2}\left((1-x)\psi_{y,N}(\alpha)\right) &= -y.\frac{\partial}{\partial x}\left(-\psi_{y,N}(\alpha)+\frac{d\psi_{y,N}}{d\alpha}(\alpha)\cdot\frac{\partial\alpha}{\partial x}\cdot(1-x)\right), \\
&= -y.\left(-\frac{d\psi_{y,N}}{d\alpha}(\alpha)\cdot\frac{\partial\alpha}{\partial x}+\frac{d\psi_{y,N}}{d\alpha}(\alpha)\frac{\partial}{\partial x}\left(\frac{\partial\alpha}{\partial x}\cdot(1-x)\right)\right. \\
&\qquad \left. + \frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha)\left(\frac{\partial\alpha}{\partial x}\right)^2(1-x)\right), \\
&= -y.\frac{d\psi_{y,N}}{d\alpha}\cdot\left(\frac{\partial}{\partial x}\left(\frac{\partial\alpha}{\partial x}\cdot(1-x)\right)-\frac{\partial\alpha}{\partial x}\right) \\
&\qquad - y\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha)\left(\frac{\partial\alpha}{\partial x}\right)^2(1-x), \\
&= -y\left(\frac{\partial\alpha}{\partial x}\right)^2(1-x)\cdot\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha).
\end{aligned}$$

The last step is remarkable because it uses the fact that $\frac{\partial}{\partial x}\left(\frac{\partial\alpha}{\partial x}\cdot(1-x)\right)-\frac{\partial\alpha}{\partial x}=0$. If we summarize the previous arguments we conclude that

$$\frac{\partial^2 R_N}{\partial x^2}(x,y) = -y\left(\frac{\partial\alpha}{\partial x}\right)^2(1-x)\cdot\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha),$$

and since $y>0$, $\left(\frac{\partial\alpha}{\partial x}\right)^2>0$ and $1-x>0$ it suffices for our property to show that

$$\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha) > 0, \qquad \forall\alpha\in(0,\infty)\ .$$

We are going to proof this by means of induction and the fact that

$$\psi_{y,1}(\alpha) = \frac{1-y+y\alpha}{1-y+\alpha},$$

$$\psi_{y,N+1}(\alpha) = \frac{\psi_{y,N}(\alpha)}{\psi_{y,N}(\alpha)+\alpha}.$$

We will prove first that

$$\max(1-\alpha,0)<\psi_{y,N}(\alpha)<1, \qquad \forall\alpha\in(0,\infty),$$

then we will prove that

$$-1<\frac{d\psi_{y,N}}{d\alpha}(\alpha)<0, \qquad \forall\alpha\in(0,\infty),$$

and then we have all ingredients to show that

$$\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha) > 0, \qquad \forall \alpha \in (0,\infty).$$

First we show that these three properties all hold for $\psi_{y,1}$.
Obviously, since $1-y>0$, $\alpha>0$ and $y>0$

$$\psi_{y,1}(\alpha) = \frac{1-y+y\alpha}{1-y+\alpha} > 0.$$

and also since $y<1$ and $\alpha>0$

$$\begin{aligned}
y\alpha-\alpha^2 &< y\alpha &< \alpha,\\
1-y+y\alpha-\alpha^2 &< 1-y+y\alpha &< 1-y+\alpha,\\
(1-\alpha)(1-y+\alpha) &< 1-y+y\alpha &< 1-y+\alpha,\\
1-\alpha &< \frac{1-y+y\alpha}{1-y+\alpha} &< 1,\\
1-\alpha &< \psi_{y,1}(\alpha) &< 1.
\end{aligned}$$

This proves that $\max(1-\alpha,0) < \psi_{y,1}(\alpha) < 1$.
It is easy to derive that

$$\frac{d\psi_{y,1}}{d\alpha}(\alpha) = -\left(\frac{1-y}{1-y+\alpha}\right)^2.$$

Since $1-y>0$ and $\alpha>0$ it is easy to see that

$$\begin{aligned}
0 &< 1-y &< 1-y+\alpha,\\
0 &< \frac{1-y}{1-y+\alpha} &< 1,\\
0 &< \left(\frac{1-y}{1-y+\alpha}\right)^2 &< 1,\\
-1 &< -\left(\frac{1-y}{1-y+\alpha}\right)^2 &< 0,\\
-1 &< \frac{d\psi_{y,1}}{d\alpha}(\alpha) &< 0.
\end{aligned}$$

This ends the proof of the second property for $\psi_{y,1}$.
Since $1-y>0$ and $\alpha>0$ is easy to derive that

$$\frac{d^2\psi_{y,1}}{d\alpha^2}(\alpha)=\frac{2(1-y)^2}{(1-y+\alpha)^3}>0.$$

This ends the proof for the case $N=1$ that for all $\alpha\in(0,1)$

$$\begin{aligned}\max(1-\alpha,0) &< \psi_{y,N}(\alpha) &< 1,\\ -1 &< \frac{d\psi_{y,N}}{d\alpha}(\alpha) &< 0,\\ & \frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha) &> 0.\end{aligned}$$

Now we prove by induction that this property holds for every $N\in\mathbb{N}^+$.
For this we use only the fact that

$$\psi_{y,N+1}(\alpha)=\frac{\psi_{y,N}(\alpha)}{\psi_{y,N}(\alpha)+\alpha}.$$

Suppose that for an arbitrary $N\in\mathbb{N}^+$ the property $\max(1-\alpha,0)<\psi_{y,N}(\alpha)<1$ holds then, since $\alpha>0$, obviously

$$\begin{aligned}0 &< \psi_{y,N}(\alpha) &< \psi_{y,N}(\alpha)+\alpha,\\ 0 &< \frac{\psi_{y,N}(\alpha)}{\psi_{y,N}(\alpha)+\alpha} &< 1,\\ 0 &< \psi_{y,N+1}(\alpha) &< 1,\end{aligned}$$

But also the following holds since $\psi_{y,N}(\alpha)>\max(1-\alpha,0)$ and $\alpha>0$

$$\begin{aligned}\psi_{y,N}(\alpha) &> 1-\alpha,\\ \psi_{y,N}(\alpha)-(1-\alpha) &> 0,\\ \alpha\cdot\frac{\psi_{y,N}(\alpha)-(1-\alpha)}{\psi_{y,N}(\alpha)+\alpha} &> 0,\\ 1-\alpha+\alpha\cdot\frac{\psi_{y,N}(\alpha)-(1-\alpha)}{\psi_{y,N}(\alpha)+\alpha} &> 1-\alpha,\\ \frac{\psi_{y,N}(\alpha)}{\psi_{y,N}(\alpha)+\alpha} &> 1-\alpha,\\ \psi_{y,N+1}(\alpha) &> 1-\alpha.\end{aligned}$$

which proves by means of induction that for all $N \in \mathbb{N}^+$

$$\max(1-\alpha, 0) \;<\; \psi_{y,N}(\alpha) \;<\; 1, \qquad \forall \alpha \in (0, \infty),$$

Let us now assume that for an arbitrary $N \in \mathbb{N}^+$ for all $\alpha \in (0, \infty)$ $-1 < \frac{d\psi_{y,N}}{d\alpha}(\alpha) < 0$ holds. From

$$\psi_{y,N+1}(\alpha) = \frac{\psi_{y,N}(\alpha)}{\psi_{y,N}(\alpha) + \alpha},$$

follows directly that

$$\frac{d\psi_{y,N+1}}{d\alpha}(\alpha) = \frac{\frac{d\psi_{y,N}}{d\alpha}(\alpha).\alpha - \psi_{y,N}(\alpha)}{(\psi_{y,N}(\alpha) + \alpha)^2}.$$

If we now start with our assumption and then proceed as follows

$$\begin{aligned}
-1 \;&<\; \frac{d\psi_{y,N}}{d\alpha}(\alpha) \;<\; 0, \\
-\alpha \;&<\; \frac{d\psi_{y,N}}{d\alpha}(\alpha) \cdot \alpha \;<\; 0, \\
-(\psi_{y,N}(\alpha) + \alpha) \;&<\; \frac{d\psi_{y,N}}{d\alpha}(\alpha) \cdot \alpha - \psi_{y,N}(\alpha) \;<\; -\psi_{y,N}(\alpha), \\
-\frac{1}{\psi_{y,N}(\alpha) + \alpha} \;&<\; \frac{\frac{d\psi_{y,N}}{d\alpha}(\alpha) \cdot \alpha - \psi_{y,N}(\alpha)}{(\psi_{y,N}(\alpha) + \alpha)^2} \;<\; -\frac{\psi_{y,N}(\alpha)}{(\psi_{y,N}(\alpha) + \alpha)^2}, \\
-\frac{1}{\psi_{y,N}(\alpha) + \alpha} \;&<\; \frac{d\psi_{y,N+1}}{d\alpha}(\alpha) \;<\; -\frac{\psi_{y,N}(\alpha)}{(\psi_{y,N}(\alpha) + \alpha)^2},
\end{aligned}$$

We already have proved that

$$\begin{aligned}
\psi_{y,N}(\alpha) \;&>\; 1 - \alpha, \\
\psi_{y,N}(\alpha) + \alpha \;&>\; 1, \\
\frac{1}{\psi_{y,N}(\alpha) + \alpha} \;&<\; 1, \\
-\frac{1}{\psi_{y,N}(\alpha) + \alpha} \;&>\; -1.
\end{aligned}$$

We also have proved that $\psi_{y,N}(\alpha) > 0$ and since also $\alpha > 0$ it is obvious that

$$-\frac{\psi_{y,N}(\alpha)}{(\psi_{y,N}(\alpha)+\alpha)^2} < 0.$$

If we combine the previous expressions we can now conclude that

$$-1 \;<\; \frac{d\psi_{y,N+1}}{d\alpha}(\alpha) \;<\; 0$$

This ends the proof by induction of the fact that for all $N \in I\!N^+$ and $\alpha \in (0,\infty)$ $-1 < \frac{d\psi_{y,N+1}}{d\alpha}(\alpha) < 0$.

By means of the properties of $\psi_{y,N}(\alpha)$ and $\frac{d\psi_{y,N}}{d\alpha}(\alpha)$ shown in the previous part we are now able to derive the property that we are most interested in. Again we use the argument of induction.

Assume that for an arbitrary $N \in I\!N$ the property $\frac{d^2\psi_{y,N+1}}{d\alpha^2}(\alpha) > 0$ holds for all $\alpha \in (0,1)$. We show that this implies that this property also holds for $N+1$. Our recurrent relation for $\psi_{y,N}(\alpha)$ can be written as follows:

$$\psi_{y,N+1}(\alpha).[\psi_{y,N}(\alpha)+\alpha] \;=\; \psi_{y,N}(\alpha).$$

We can differentiate the equation one time as follows

$$\frac{d\psi_{y,N+1}}{d\alpha}(\alpha).[\psi_{y,N}(\alpha)+\alpha] + \psi_{y,N+1}(\alpha).[\frac{d\psi_{y,N}}{d\alpha}(\alpha)+1] \;=\; \frac{d\psi_{y,N}}{d\alpha}(\alpha).$$

If we differentiate this equation again we obtain

$$\frac{d^2\psi_{y,N+1}}{d\alpha^2}(\alpha).[\psi_{y,N}(\alpha)+\alpha] \;+\; 2\frac{d\psi_{y,N+1}}{d\alpha}.[\frac{d\psi_{y,N}}{d\alpha}(\alpha)+1] + \psi_{y,N+1}(\alpha).\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha) \;=\; \frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha).$$

We rewrite the last equation and get that

$$\frac{d^2\psi_{y,N+1}}{d\alpha^2}(\alpha) = \frac{\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha).[1-\psi_{y,N+1}(\alpha)] - 2\frac{d\psi_{y,N+1}}{d\alpha}.[\frac{d\psi_{y,N}}{d\alpha}(\alpha)+1]}{\psi_{y,N}(\alpha)+\alpha}.$$

Now we will use all properties again.
First we assume that

$$\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha) > 0,$$

next we have shown that

$$\psi_{y,N+1}(\alpha) < 1,$$

$$1 - \psi_{y,N+1}(\alpha) > 0,$$

which makes

$$\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha).[1 - \psi_{y,N+1}(\alpha)] > 0.$$

We have also shown that

$$\frac{d\psi_{y,N+1}}{d\alpha} < 0,$$

$$-2 \cdot \frac{d\psi_{y,N+1}}{d\alpha} > 0,$$

and that

$$\frac{d\psi_{y,N}}{d\alpha}(\alpha) > -1,$$

$$\frac{d\psi_{y,N}}{d\alpha}(\alpha) + 1 > 0,$$

which also makes

$$-2 \cdot \frac{d\psi_{y,N+1}}{d\alpha}.[\frac{d\psi_{y,N}}{d\alpha}(\alpha) + 1] > 0.$$

If we combine this we get that the numerator is positive:

$$\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha).[1 - \psi_{y,N+1}(\alpha)] - 2\frac{d\psi_{y,N+1}}{d\alpha}.[\frac{d\psi_{y,N}}{d\alpha}(\alpha) + 1] > 0.$$

Because we have shown that $\psi_{y,N}(\alpha) > 0$ and $\alpha > 0$ we conclude that the denominator is also positive:

$$\psi_{y,N}(\alpha) + \alpha > 0,$$

And thus we have shown that

$$\frac{d^2\psi_{y,N+1}}{d\alpha^2}(\alpha) = \frac{\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha).[1 - \psi_{y,N+1}(\alpha)] - 2\frac{d\psi_{y,N+1}}{d\alpha}.[\frac{d\psi_{y,N}}{d\alpha}(\alpha) + 1]}{\psi_{y,N}(\alpha) + \alpha} > 0.$$

By induction we now have proved that for all $N \in \mathbb{N}^+$ the following holds

$$\frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha) > 0, \qquad \forall \alpha \in (0,\infty).$$

We remind the reader that we have already derived that

$$\frac{\partial^2 R_N}{\partial x^2}(x,y) \;=\; -y\left(\frac{\partial \alpha}{\partial x}\right)^2 (1-x)\cdot \frac{d^2\psi_{y,N}}{d\alpha^2}(\alpha),$$

And therefore we have also shown that for all $N \in \mathbb{N}^+$ and for all $(x,y) \in (0,1)^2$

$$\frac{\partial^2 R_N}{\partial x^2}(x,y) < 0.$$

Because of the symmetry of R_N described in "R_N property 4" it is obvious that the following is also true for all $N \in \mathbb{N}^+$ and for all $(x,y) \in (0,1)^2$:

$$\frac{\partial^2 R_N}{\partial y^2}(x,y) < 0.$$

This completes the proof of "R_N property 9".

■

Theorem 15 (R_N **property 10)** *For all $N \in \mathbb{N}^+$ and $(x,y) \in (0,1)^2$ the following holds:*

$$\| \;\min(x,y) - R_N(x,y)\; \| \;<\; \frac{1}{4N}. \tag{A.15}$$

Proof. First we define

$$\begin{aligned} D_N \;&:\; (0,1)^2 \;\mapsto\; (0,1), \\ D_N(x,y) \;&\stackrel{\text{def}}{=}\; \min(x,y) - R_N(x,y). \end{aligned} \tag{A.16}$$

Then it is obvious from "R_N property 4" that functions D_N are symmetric:

$$D_N(x,y) = D_N(y,x), \qquad \forall (x,y) \in (0,1)^2.$$

"R_N property 5" combined with the symmetry of R_N implies that

$$0 \;<\; D_N(x,y) \;<\; 1, \qquad \forall (x,y) \in (0,1)^2.$$

Because of the symmetry of D_N and because D_N is strictly positive it is sufficient for the proof to show that for all $N \in \mathbb{N}^+$

$$D_N(x,y) \;<\; \frac{1}{4N}, \quad \forall (x,y) \in (0,1)^2 \;\text{ with }\; x \le y.$$

Therefore we consider a pair $(x, y) \in (0, 1)^2$ with $x \leq y$. In this case the function D_N reduces to

$$D_N(x, y) \quad = \quad x - R_N(x, y),$$

and from the recursive equation of "R_N property 3" we derive that

$$\begin{aligned}
D_1(x, y) &= \frac{x^2(1-y)}{x + y(1-x)}, \\
D_{N+1}(x, y) &= x - R_{N+1}(x, y), \\
&= x - x \cdot y \cdot \frac{1 - R_N(x, y)}{x + y - xy - R_N(x, y)}, \\
&= x - x \cdot y \cdot \frac{1 - x + [x - R_N(x, y)]}{x + y - xy - x + [x - R_N(x, y)]}, \\
&= x - x \cdot y \cdot \frac{1 - x + D_N(x, y)}{y(1-x) + D_N(x, y)}, \\
&= \frac{x(1-y)D_N(x, y)}{y(1-x) + D_N(x, y)}.
\end{aligned}$$

Since $0 < x \leq y$ we derive for all $z \in \mathbb{R},\ z > 0$ that

$$\begin{aligned}
x &\leq y, \\
-y &\leq -x, \\
-xy &\leq -x^2, \\
x - xy &\leq x - x^2, \\
x(1-y) &\leq x(1-x) \leq \frac{1}{4}, \\
4x(1-y) &\leq 1, \\
4x(1-y)z &\leq z, \\
x + 4x(1-y)z &\leq y + z,
\end{aligned}$$

$$\begin{aligned} x - xy + 4x(1-y)z &\leq y - xy + z, \\ x(1-y) + 4x(1-y)z &\leq y(1-x) + z, \\ x(1-y)(1+4z) &\leq y(1-x) + z, \\ \frac{x(1-y)}{y(1-x)+z} &\leq \frac{1}{1+4z}, \\ \frac{x(1-y)z}{y(1-x)+z} &\leq \frac{z}{1+4z}. \end{aligned}$$

We can easily derive that $\frac{z}{1+4z}$ is a monotonically increasing function of z. Therefore let again $(x,y) \in (0,1)^2$, $z_1 \in \mathbb{R}$ and $z_2 \in \mathbb{R}$ such that $z_1 < z_2$, then

$$\begin{aligned} z_1 &< z_2, \\ z_1 + 4z_1z_2 &< z_2 + 4z_1z_2, \\ z_1[1+4z_2] &< z_2[1+4z_1], \\ \frac{z_1}{1+4z_1} &< \frac{z_2}{1+4z_2}. \end{aligned}$$

We can now prove by induction that $D_N(x,y) < \frac{1}{4N}$.
For $N = 1$ we have

$$D_1(x,y) = \frac{x(1-y).x}{y(1-x)+x} \leq \frac{x}{1+4x} < \frac{x}{4x} = \frac{1}{4} = \frac{1}{4N}.$$

If we assume for an arbitrary $N \in \mathbb{N}^+$ that $D_N(x,y) < \frac{1}{4N}$ then we derive by monotonicity of $\frac{z}{1+4z}$ that

$$D_{N+1}(x,y) = \frac{x(1-y)D_N(x,y)}{y(1-x)+D_N(x,y)} \leq \frac{D_N(x,y)}{1+4D_N(x,y)} < \frac{\frac{1}{4N}}{1+4\cdot\frac{1}{4N}} = \frac{1}{4(N+1)}.$$

By the induction argument we have now shown that for all $N \in \mathbb{N}^+$ and for all $(x,y) \in (0,1)^2$ with $x \leq y$

$$D_N(x,y) < \frac{1}{4N}.$$

This ends the proof of "R_N property 10".

■

Theorem 16 **(R_N property 11)** *For all $N \in \mathbb{N}^+$ and $(x,y) \in (0,1)^2$ the following holds:*

$$\begin{aligned} x < y \quad &\Rightarrow \quad \frac{\partial R_N}{\partial x}(x,y) > \frac{\partial R_N}{\partial y}(x,y), \\ x > y \quad &\Rightarrow \quad \frac{\partial R_N}{\partial x}(x,y) < \frac{\partial R_N}{\partial y}(x,y). \end{aligned} \tag{A.17}$$

Proof. Again we will use induction to proof this property.
Therefore we first show the property for R_N with $N = 1$. From

$$R_1(x,y) = \frac{xy}{x+y-xy},$$

follows directly that

$$\frac{\partial R_1}{\partial x}(x,y) = \frac{y^2}{(x+y-xy)^2}.$$

From the symmetry between arguments x and y in $R_1(x,y)$ it is clear that then also

$$\frac{\partial R_1}{\partial y}(x,y) = \frac{x^2}{(x+y-xy)^2}.$$

And therefore we can conclude that

$$\begin{aligned} \frac{\partial R_1}{\partial x}(x,y) - \frac{\partial R_1}{\partial y}(x,y) &= \frac{y^2 - x^2}{(x+y-xy)^2}, \\ &= (y-x) \cdot \frac{x+y}{(x+y-xy)^2}, \end{aligned}$$

and since obviously for all $(x,y) \in (0,1)^2$

$$\frac{x+y}{(x+y-xy)^2} > 0,$$

the property is shown for the case $N = 1$.

From "R_N property 3" we know that

$$R_{N+1}(x,y) = x \cdot y \cdot \frac{1 - R_N(x,y)}{x+y-xy-R_N(x,y)}.$$

From this property we can derive that

$$\frac{\partial R_{N+1}}{\partial x}(x,y) = \frac{y(1-R_N(x,y))(y-R_N(x,y)) + xy(1-x)(1-y)\frac{\partial R_N}{\partial x}(x,y)}{(x+y-xy)^2}.$$

Because of the symmetry property denoted by "R_N property 4" we can conclude that then also

$$\frac{\partial R_{N+1}}{\partial y}(x,y) = \frac{x(1-R_N(x,y))(x-R_N(x,y)) + xy(1-x)(1-y)\frac{\partial R_N}{\partial y}(x,y)}{(x+y-xy)^2}.$$

If we combine these two expressions we get

$$\begin{aligned}\frac{\partial R_{N+1}}{\partial x}(x,y) - \frac{\partial R_{N+1}}{\partial y}(x,y) &= (y-x)\cdot\frac{(1-R_N(x,y))(x+y-R_N(x,y))}{(x+y-xy)^2} + \ldots \\ &\quad + \left(\frac{\partial R_N}{\partial x}(x,y) - \frac{\partial R_N}{\partial y}(x,y)\right)\cdot\frac{xy(1-x)(1-y)}{(x+y-xy)^2}.\end{aligned} \tag{A.18}$$

It is easy to see from R_N properties 3 and 4 that for all $(x,y) \in (0,1)^2$, $1 - R_N(x,y) > 0$ and $x + y - R_N(x,y) > 0$ and that therefore for all $(x,y) \in (0,1)^2$

$$\frac{(1-R_N(x,y))(x+y-R_N(x,y))}{(x+y-xy)^2} > 0.$$

It is also obvious that for all $(x,y) \in (0,1)^2$

$$\frac{xy(1-x)(1-y)}{(x+y-xy)^2} > 0,$$

Let us assume now that the property that we want to prove holds for an arbitrary $N \in \mathbb{N}^+$.

Then first let $(x,y) \in (0,1)^2$ such that $x < y$.
Then $y - x > 0$ and by assumption

$$\frac{\partial R_N}{\partial x}(x,y) - \frac{\partial R_N}{\partial y}(x,y) > 0.$$

This combination in Expression (A.18) makes that

$$\frac{\partial R_{N+1}}{\partial x}(x,y) - \frac{\partial R_{N+1}}{\partial y}(x,y) > 0.$$

This shows the first part of the property for $N + 1$.

Finally let $(x, y) \in (0, 1)^2$ such that $x > y$.
Then $y - x < 0$ and by assumption

$$\frac{\partial R_N}{\partial x}(x, y) - \frac{\partial R_N}{\partial y}(x, y) \quad < \quad 0.$$

This combination in Expression (A.18) makes that

$$\frac{\partial R_{N+1}}{\partial x}(x, y) - \frac{\partial R_{N+1}}{\partial y}(x, y) \quad < \quad 0.$$

This shows the second part of the property for $N + 1$.

We conclude that the property also holds for $N + 1$.
By the induction argument we conclude that our property holds for every $N \in \mathbb{N}^+$.
This ends the proof of "R_N property 11".

■

Theorem 17 **(R_N property 12)** *For all $y \in [0, 1]$ and for all $N \in \mathbb{N}^+$*

$$\lim_{x \downarrow 0} \frac{R_N(x, y)}{x} = 1.$$

Proof. Let $N \in \mathbb{N}^+$, then according to "R_N property 3"

$$\lim_{x \downarrow 0} \frac{R_N(x, y)}{x} \quad = \quad \lim_{x \downarrow 0} \quad y \cdot \frac{1 - R_{N-1}(x, y)}{x + y - xy - R_{N-1}(x, y)}.$$

The functions R_{N-1} are continuous on $[0, 1]^2$ for all $N \in \mathbb{N}^+$, and

$$\lim_{x \downarrow 0} \quad R_{N-1}(x, y) = 0.$$

Therefore

$$\begin{aligned} \lim_{x \downarrow 0} \quad y \cdot \frac{1 - R_{N-1}(x, y)}{x + y - xy - R_{N-1}(x, y)} \quad &= \quad y \cdot \frac{1 - 0}{0 + y - 0.y - 0}, \\ &= \quad y \cdot \frac{1}{y}, \\ &= \quad 1. \end{aligned}$$

This ends the proof of "R_N property 12".

■

A.3 The Function $\Omega_{c,N}$ and Its Properties

The function Ω_N is introduced in Expression (2.49) and is used throughout Chapter 2 to describe the equations that link neighbouring two-machine-systems.

Definition 3 *Let $c \in \mathbb{R}$, $c \in (0,1]$ and $N \in \mathbb{N}^+$, then*

$$\begin{aligned} \Omega_{c,N} &: [0,1] \times [0,1] \mapsto [0,c], \\ \Omega_{c,N}(x,y) &\stackrel{\text{def}}{=} c.(1 - y.\Psi_N(x,y)). \end{aligned} \tag{A.19}$$

Theorem 18 ($\Omega_{c,N}$ **property 1)** *For all $(x,y) \in (0,1)^2$, for all $c \in (0,1]$ and for all $N \in \mathbb{N}^+$*

$$\Omega_{c,N}(x,y) = c \cdot \frac{R_N(1-x,1-y)}{1-x}. \tag{A.20}$$

Proof. This is a direct consequence of the definitions of R_N and $\Omega_{c,N}$.

■

Theorem 19 ($\Omega_{c,N}$ **property 2)** *$\Omega_{c,N}$ and all its partial derivatives of any order exist and are continuous on $(0,1)^2$.*

Proof. This is a direct consequence of "Ψ_N property 4".

■

Theorem 20 ($\Omega_{c,N}$ **property 3)** *For all $(x,y) \in (0,1)^2$, for all $c \in (0,1]$ and for all $N \in \mathbb{N}^+$*

$$\begin{aligned} \frac{\partial \Omega_{c,N}}{\partial x}(x,y) &> 0, \\ \frac{\partial \Omega_{c,N}}{\partial y}(x,y) &< 0. \end{aligned} \tag{A.21}$$

Proof. From "$\Omega_{c,N}$ property 1" follows that

$$\Omega_{c,N}(x,y) = c \cdot \frac{R_N(1-x,1-y)}{1-x}.$$

If we differentiate both sides of this expression to the variable x we obtain

$$\frac{\partial \Omega_{c,N}}{\partial x}(x,y) = c \cdot \frac{-(1-x) \cdot \frac{\partial R_N}{\partial x}(1-x,1-y) + R_N(1-x,1-y)}{(1-x)^2}.$$

From "R_N property 8" with $(1-x,1-y)$ substituted for (x,y) follows

$$\begin{aligned} \frac{\partial R_N}{\partial x}(1-x,1-y) &< \frac{R_N(1-x,1-y)}{1-x}, \\ -(1-x) \cdot \frac{\partial R_N}{\partial x}(1-x,1-y) + R_N(1-x,1-y) &> 0. \end{aligned}$$

This combined with the fact that $(1-x)^2 > 0$ makes that for all $(x,y) \in (0,1)^2$, for all $c \in (0,1]$ and for all $N \in \mathbb{N}^+$

$$\frac{\partial \Omega_{c,N}}{\partial x}(x,y) > 0.$$

From "$\Omega_{c,N}$ property 1"

$$\Omega_{c,N}(x,y) = c \cdot \frac{R_N(1-x,1-y)}{1-x},$$

follows that

$$\frac{\partial \Omega_{c,N}}{\partial y}(x,y) = -\frac{c}{1-x} \cdot \frac{\partial R_N}{\partial y}(1-x,1-y).$$

From "R_N property 8" with $(1-x,1-y)$ substituted for (x,y) follows

$$\frac{\partial R_N}{\partial y}(1-x,1-y) > 0,$$

and therefore we have shown that for all $(x,y) \in (0,1)^2$, for all $c \in (0,1]$ and for all $N \in \mathbb{N}^+$

$$\frac{\partial \Omega_{c,N}}{\partial y}(x,y) < 0.$$

This ends the proof of "$\Omega_{c,N}$ property 3".

■

A.4 The Function Q_N and Its Properties

The function Q_N is used only in the proof of Theorem 30 in Appendix B. For the definition of functions Q_N for $N \in \mathbb{N}$ we first define the sets U and V as follows:

Definition 4

$$\begin{aligned} U &\stackrel{\text{def}}{=} \{\ (x,y) \in (0,1) \times (0,1)\ \mid x > y\ \}, \\ V &\stackrel{\text{def}}{=} \{\ (x,y) \in [0,1] \times (0,1]\ \mid x \geq y\ \}. \end{aligned} \tag{A.22}$$

By means of this definition we define the functions Q_N for $N \in \mathbb{N}$ *implicitly* by

Definition 5

$$\begin{aligned} Q_N &: V \mapsto [0,1]\,, \\ R_N(x, Q_N(x,y)) &= y, \qquad \forall\ (x,y) \in V. \end{aligned} \tag{A.23}$$

The *implicit* definition of functions Q_N is consistent since it can easily be shown by means of the properties of the functions R_N that,

Lemma 1 *For all* $(a,b) \in V$ *the following equation has a* unique *solution for* x *in* $[0,1]$*:*

$$R_N(a,x) = b.$$

Proof. In order to show that such a function Q_N exists uniquely we show that for two given constants $a \in (0,1]$ and $b \in [0,a]$ there exists a unique solution for $x \in [0,1]$ in the equation:

$$R_N(a,x) = b. \tag{A.24}$$

To show that there exists a solution we construct the function $q_{N,a,b}(x)$:

$$\begin{aligned} q_{N,a,b} \quad &: \quad [0,1] \;\mapsto\; [-b, a-b], \\ q_{N,a,b}(x) \quad &\stackrel{\text{def}}{=} \quad R_N(a,x) - b. \end{aligned} \tag{A.25}$$

Then obviously:

$$\begin{aligned} q_{N,a,b} \quad &\text{is} \quad \text{continuous}, \\ q_{N,a,b}\,(0) \;&=\; -b \;\leq\; 0, \\ q_{N,a,b}\,(1) \;&=\; a-b \;\geq\; 0. \end{aligned} \tag{A.26}$$

This combination makes that there exists a root $x \in [0,1]$ of function $q_{N,a,b}(x)$. From the properties of R_N in Appendix A it is easy to see that the derivative of $q_{N,a,b}(x)$ exists for all $x \in (0,1)$. From the fact that $\frac{\partial R_N}{\partial x}(x,y) > 0$ from "R_N property 8" in Appendix A follows that if $a \in (0,1)$ then $q'_{N,a,b}(x) > 0$ for all $x \in (0,1)$. As a direct consequence of the mean value theorem we can conclude that, if $a \in (0,1]$ and $b \in [0,a]$ then $q_{N,a,b}(x)$ has a unique root, and therefore there exists a unique solution $x \in [0,1]$ for Equation (A.24). We now have shown that there exists a unique function Q_N that has the following properties:

$$\begin{aligned} Q_N \quad &: \quad V \;\mapsto\; [0,1]\,, \\ R_N(x, Q_N(x,y)) \;&=\; y, \qquad \forall\,(x,y) \in V. \end{aligned} \tag{A.27}$$

■

Theorem 21 (Q_N **property 1)** *For all* $x \in (0,1)$ *and for all* $N \in \mathbb{N}^+$

$$\begin{aligned} Q_N(x,0) \;&=\; 0, \\ Q_N(x,x) \;&=\; 1, \\ Q_N(1,x) \;&=\; x. \end{aligned} \tag{A.28}$$

Proof. This is obvious if we combine the definition of Q_N with

$$\begin{aligned} R_N(x,0) &= 0, \\ R_N(x,1) &= x, \\ R_N(1,x) &= x. \end{aligned}$$

This ends the proof of "Q_N property 1". ∎

Theorem 22 **(Q_N property 2)** *Q_N and all its partial derivatives of any order exist and are continuous on U.*

Theorem 23 **(Q_N property 3)** *For all $(x,y) \in U$ and for all $N \in \mathbb{N}^+$*

$$\begin{aligned} \frac{\partial Q_N}{\partial x}(x,y) &= -\frac{\frac{\partial R_N}{\partial x}(x, Q_N(x,y))}{\frac{\partial R_N}{\partial y}(x, Q_N(x,y))} < 0, \\ \frac{\partial Q_N}{\partial y}(x,y) &= \frac{1}{\frac{\partial R_N}{\partial y}(x, Q_N(x,y))} > 0. \end{aligned} \tag{A.29}$$

Proof. The definition of functions Q_N states that for all $(x,y) \in U$ and for all $N \in \mathbb{N}^+$

$$R_N(x, Q_N(x,y)) = y.$$

If we differentiate this equation to the variable x we obtain

$$\frac{\partial R_N}{\partial x}(x, Q_N(x,y)) + \frac{\partial R_N}{\partial y}(x, Q_N(x,y)) \cdot \frac{\partial Q_N}{\partial x}(x,y) = 0,$$

which leads directly to the first equation.
If we differentiate the equation to the variable y we obtain

$$\frac{\partial R_N}{\partial y}(x, Q_N(x,y)) \cdot \frac{\partial Q_N}{\partial y}(x,y) = 1,$$

which leads directly to the second equation.
From R_N properties 8 and 4 follows that

$$\begin{aligned} \frac{\partial R_N}{\partial x}(x, Q_N(x,y)) &> 0, \\ \frac{\partial R_N}{\partial y}(x, Q_N(x,y)) &> 0, \end{aligned}$$

which leads to the inequalities in this property.
This ends the proof of "Q_N property 3".

■

Theorem 24 **(Q_N property 4)** *For all $(x,y) \in U$ and for all $N \in \mathbb{N}^+$*

$$Q_N(x,y) \quad > \quad y. \tag{A.30}$$

Proof. Let $(x,y) \in U$.
From a combination of R_N properties 4 and 5 follows that for all $(x,y) \in (0,1)^2$ and for all $N \in \mathbb{N}^+$

$$R_N(x,y) \qquad < \qquad y.$$

Since $Q_N(x,y) \in (0,1)$ on U we can substitute $Q_N(x,y)$ for y in this inequality and we get

$$R_N(x,Q_N(x,y)) \qquad < \qquad Q_N(x,y),$$

By the definition of $Q_N(x,y)$ this reduces immediately to

$$y \qquad < \qquad Q_N(x,y).$$

This ends the proof of "Q_N property 4".

■

Theorem 25 **(Q_N property 5)** *For all $(x,y) \in U$ and for all $N \in \mathbb{N}^+$*

$$\frac{\partial Q_N}{\partial y}(x,y) \quad > \quad \frac{Q_N(x,y)}{y}. \tag{A.31}$$

Proof. A combination of R_N properties 4 and 8 makes that for all $(x,y) \in (0,1)^2$ and for all $N \in \mathbb{N}^+$

$$\frac{\partial R_N}{\partial y}(x,y) \quad < \quad \frac{R_N(x,y)}{y}.$$

This means, since $Q_N(x,y) \in (0,1)$ on U, that we may substitute $Q_N(x,y)$ for y and get

$$\frac{\partial R_N}{\partial y}(x,Q_N(x,y)) \quad < \quad \frac{R_N(x,Q_N(x,y))}{Q_N(x,y)}.$$

Since $\frac{\partial R_N}{\partial y}(x,Q_N(x,y)) > 0$, $R_N(x,Q_N(x,y)) > 0$ and $y > 0$ we can rewrite this as follows

$$\frac{1}{\frac{\partial R_N}{\partial y}(x,Q_N(x,y))} \quad > \quad \frac{Q_N(x,y)}{R_N(x,Q_N(x,y))}.$$

Because of the definition of $Q_N(x,y)$), and because of "Q_N property 3" the second part we deduce that

$$\frac{\partial Q_N}{\partial y}(x,y) \quad > \quad \frac{Q_N(x,y)}{y}.$$

This completes the proof of "Q_N property 5".

■

Theorem 26 (Q_N **property 6**) *For all* $x \in (0,1)$ *and for all* $N \in \mathbb{N}^+$

$$\lim_{y \downarrow 0} \quad \frac{y}{Q_N(x,y)} = 1.$$

Proof. We start with a combination of "R_N property 4" and "R_N property 12". For all $x \in [0,1]$ and for all $N \in \mathbb{N}^+$

$$\lim_{a \downarrow 0} \quad \frac{R_N(x,a)}{a} = 1.$$

We can substitute $Q_N(x,y)$ for a

$$\lim_{Q_N(x,y) \downarrow 0} \quad \frac{R_N(x, Q_N(x,y))}{Q_N(x,y)} = 1.$$

Since for all $x \in (0,1]$, $Q_N(x,0) = 0$ and since Q_N is continuous on V and therefore certainly continuous in points $(x,0)$ we conclude that for some function f

$$\lim_{Q_N(x,y) \downarrow 0} \quad f(Q_N(x,y)) \quad = \quad \lim_{y \downarrow 0} \quad f(Q_N(x,y)).$$

If we apply this on our equation we conclude that

$$\lim_{y \downarrow 0} \quad \frac{R_N(x, Q_N(x,y))}{Q_N(x,y)} = 1.$$

By the definition of Q_N this reduces to

$$\lim_{y \downarrow 0} \quad \frac{y}{Q_N(x,y)} = 1.$$

This completes the proof of "Q_N property 6".

■

Appendix B

Proofs for the Forward and Backward Aggregation Method in Chapter 2

In this appendix we will state 4 theorems with their proofs. The theorems concern the properties of the forward and backward aggregation algorithm that has been developed in Chapter 2. The forward and backward aggregation algorithm is fully described in Expression (2.51).

First we will give a some definitions and two lemmas by means of which we reformulate the algorithm in a setting of multi-dimensional functions. In this setting it is easier to formulate and proof the theorems. That is what we will do next.

B.1 Some Definitions

We will start by collecting corresponding variables, and putting them in the shape of vectors:

Definition 6

$$\boldsymbol{p} \stackrel{\text{def}}{=} \begin{pmatrix} p_1 \\ p_2 \\ \cdot \\ \cdot \\ p_m \end{pmatrix}, \boldsymbol{p}^f(s) \stackrel{\text{def}}{=} \begin{pmatrix} p_1^f(s) \\ p_2^f(s) \\ \cdot \\ \cdot \\ p_m^f(s) \end{pmatrix}, \boldsymbol{p}^b(s) \stackrel{\text{def}}{=} \begin{pmatrix} p_1^b(s) \\ p_2^b(s) \\ \cdot \\ \cdot \\ p_m^b(s) \end{pmatrix}, \boldsymbol{N} \stackrel{\text{def}}{=} \begin{pmatrix} N_2 \\ N_3 \\ \cdot \\ \cdot \\ N_m \end{pmatrix}. \tag{B.1}$$

Then we define the vector function:

Definition 7

$$\boldsymbol{B_{p,N}} : [0,1]^m \mapsto [0,1]^m,$$

$$\boldsymbol{B_{p,N}} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_{m-1} \\ x_m \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} \Omega_{p_1,N_2}(1-x_1, 1-\Omega_{p_2,N_3}(\ldots(1-x_{m-2}, 1-\Omega_{p_{m-1},N_m}(1-x_{m-1}, 1-p_m)\ldots))) \\ \Omega_{p_2,N_3}(\ldots(1-x_{m-2}, 1-\Omega_{p_{m-1},N_m}(1-x_{m-1}, 1-p_m)\ldots)) \\ \cdot \\ \cdot \\ \Omega_{p_{m-1},N_m}(1-x_{m-1}, 1-p_m) \\ p_m \end{pmatrix}.$$

And we define the vector function:

Definition 8

$$\boldsymbol{F_{p,N}} : [0,1]^m \mapsto [0,1]^m,$$

$$\boldsymbol{F_{p,N}} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_{m-1} \\ x_m \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} p_1 \\ \Omega_{p_2,N_2}(1-x_2, 1-p_1) \\ \cdot \\ \cdot \\ \Omega_{p_{m-1},N_{m-1}}(\ldots(1-x_3, 1-\Omega_{p_2,N_2}(1-x_2, 1-p_1)\ldots)) \\ \Omega_{p_m,N_m}(1-x_m, 1-\Omega_{p_{m-1},N_{m-1}}(\ldots(1-x_3, 1-\Omega_{p_2,N_2}(1-x_2, 1-p_1)\ldots))) \end{pmatrix}.$$

We can compound the previous vector functions to get:

Definition 9

$$\begin{aligned} \boldsymbol{C_{p,N}} &: [0,1]^m \mapsto [0,1]^m, \\ \boldsymbol{C_{p,N}} &\stackrel{\text{def}}{=} \boldsymbol{F_{p,N}} \circ \boldsymbol{B_{p,N}}. \end{aligned} \tag{B.2}$$

By means of these vector functions it is possible to formulate the algorithm as follows:

$$\begin{aligned} &\textbf{Initial condition}: \\ &\boldsymbol{p}^f(0) = \boldsymbol{0}, \\ &\textbf{Backward aggregation}: \\ &\boldsymbol{p}^b(s) = \boldsymbol{B_{p,N}}(\boldsymbol{p}^f(s)), \\ &\textbf{Forward aggregation}: \\ &\boldsymbol{p}^f(s+1) = \boldsymbol{F_{p,N}}(\boldsymbol{p}^b(s)). \end{aligned} \tag{B.3}$$

Or as follows:

$$\textbf{Initial condition}:$$

$$\boldsymbol{p}^f(0) = \mathbf{0}, \tag{B.4}$$

$$\textbf{Iteration}:$$

$$\boldsymbol{p}^f(s+1) = \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{N}}(\,\boldsymbol{p}^f(s)).$$

Therefore we will examine the vector functions defined in (B.2), (B.2) and (B.2):

Lemma 2

$$\text{if } \boldsymbol{p} \in (0,1)^m \text{ then } \boldsymbol{B}_{\boldsymbol{p},\boldsymbol{N}},\ \boldsymbol{F}_{\boldsymbol{p},\boldsymbol{N}} \text{ and } \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{N}} \text{ are differentiable, } \forall\, \boldsymbol{x} \in (0,1)^m. \tag{B.5}$$

Proof. This follows directly from the following properties of the functions $\Omega_{c,N}$ for all $c \in (0,1)$ and $N \in I\!N^+$:

$$\begin{gathered}\Omega_{c,N}(\,(0,1)\times(0,1)\,) \subset (0,1),\\ \frac{\partial \Omega_{c,N}}{\partial x}(x,y) \text{ and } \frac{\partial \Omega_{c,N}}{\partial y}(x,y) \text{ both exist, } \forall (x,y) \in (0,1)\times(0,1).\end{gathered} \tag{B.6}$$

■

Therefore we can define the derivatives of the vector functions that we have just defined:

Definition 10 *For* $\boldsymbol{N} \in (I\!N^+)^{m-1}$ *and* $\boldsymbol{p} \in (0.1)^m$*:*

$$\begin{aligned}\{b(\boldsymbol{x})_{i,j}\} &\stackrel{\text{def}}{=} \boldsymbol{B}'_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{x}),\\ \{f(\boldsymbol{x})_{i,j}\} &\stackrel{\text{def}}{=} \boldsymbol{F}'_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{x}),\\ \{c(\boldsymbol{x})_{i,j}\} &\stackrel{\text{def}}{=} \boldsymbol{C}'_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{x}).\end{aligned} \tag{B.7}$$

B.2 A Lemma

Lemma 3

$$\begin{aligned} b(\boldsymbol{x})_{i,j} &\le 0, \quad \forall\, \boldsymbol{x} \in (0,1)^m, \ \forall\, (i,j) \in \{1,2,..,m\}^2,\\ f(\boldsymbol{x})_{i,j} &\le 0, \quad \forall\, \boldsymbol{x} \in (0,1)^m, \ \forall\, (i,j) \in \{1,2,..,m\}^2,\\ c(\boldsymbol{x})_{i,j} &\ge 0, \quad \forall\, \boldsymbol{x} \in (0,1)^m, \ \forall\, (i,j) \in \{1,2,..,m\}^2, \end{aligned} \tag{B.8}$$

Proof. First we refer to the Appendix A in "$\Omega_{c,N}$ property 3" where we have shown that the functions $\Omega_{c,N}$ have the following properties for all $c \in (0,1)$ and $N \in I\!N^+$:

$$\begin{aligned} \frac{\partial \Omega_{c,N}}{\partial x}(x,y) &\ge 0, \quad \forall\, (x,y) \in (0,1)\times(0,1),\\ \frac{\partial \Omega_{c,N}}{\partial y}(x,y) &\le 0, \quad \forall\, (x,y) \in (0,1)\times(0,1). \end{aligned} \tag{B.9}$$

If we examine closely the structure of the vector functions $\boldsymbol{F_{p,N}}$ and $\boldsymbol{B_{p,N}}$, then we conclude that:

$$\begin{array}{lllll} b_{i,j}(\boldsymbol{x}) \leq 0, & \forall j \in \{1,2,..,m\} & \Rightarrow & b_{i-1,j}(\boldsymbol{x}) \leq 0, & \forall j \in \{1,2,..,m\}, \\ f_{i,j}(\boldsymbol{x}) \leq 0, & \forall j \in \{1,2,..,m\} & \Rightarrow & f_{i+1,j}(\boldsymbol{x}) \leq 0, & \forall j \in \{1,2,..,m\}, \end{array} \tag{B.10}$$

We will show this in the following.
In order to show statement (B.10) we first observe that:

$$\begin{array}{rcl} \left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_{i-1} & = & \Omega_{p_{i-1},N_i}\ (1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i), \\ \left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_{i+1} & = & \Omega_{p_{i+1},N_{i+1}}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i). \end{array} \tag{B.11}$$

If we now take the partial dirivatives and differentiate both sides of Expressions (B.11) to the variable $x_j,\ j \in \{1,2,..,m\}$ we get the following:

$$\begin{array}{rcl} b_{i-1,j}(\boldsymbol{x}) & = & \frac{\partial \Omega_{p_{i-1},N_i}}{\partial x}(1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i) \cdot \frac{\partial}{\partial x_j}(1-x_{i-1}) + .. \\ & & .. + \frac{\partial \Omega_{p_{i-1},N_i}}{\partial y}(1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i) \cdot \frac{\partial}{\partial x_j}(1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i), \\ f_{i+1,j}(\boldsymbol{x}) & = & \frac{\partial \Omega_{p_{i+1},N_{i+1}}}{\partial x}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i) \cdot \frac{\partial}{\partial x_j}(1-x_{i+1}) + .. \\ & & .. + \frac{\partial \Omega_{p_{i+1},N_{i+1}}}{\partial y}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i) \cdot \frac{\partial}{\partial x_j}(1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i). \end{array} \tag{B.12}$$

This is equivalent to:

$$\begin{array}{rcl} b_{i-1,j}(\boldsymbol{x}) & = & -\frac{\partial x_{i-1}}{\partial x_j} \cdot \frac{\partial \Omega_{p_{i-1},N_i}}{\partial x}(1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i) + .. \\ & & .. - b_{i,j}(\boldsymbol{x}) \cdot \frac{\partial \Omega_{p_{i-1},N_i}}{\partial y}(1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i), \\ f_{i+1,j}(\boldsymbol{x}) & = & -\frac{\partial x_{i+1}}{\partial x_j} \cdot \frac{\partial \Omega_{p_{i+1},N_{i+1}}}{\partial x}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i) + .. \\ & & .. - f_{i,j}(\boldsymbol{x}) \cdot \frac{\partial \Omega_{p_{i+1},N_{i+1}}}{\partial y}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i). \end{array} \tag{B.13}$$

Of course

$$\begin{array}{rcl} \frac{\partial x_{i-1}}{\partial x_j} & \geq & 0, \\ \frac{\partial x_{i+1}}{\partial x_j} & \geq & 0. \end{array} \tag{B.14}$$

From (B.9) it follows that

$$\begin{array}{rcl} \frac{\partial \Omega_{p_{i-1},N_i}}{\partial x}(1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i) & \geq & 0, \\ \frac{\partial \Omega_{p_{i-1},N_i}}{\partial y}(1-x_{i-1}, 1-\left\{\boldsymbol{B_{p,N}}(\boldsymbol{x})\right\}_i) & \leq & 0, \\ \frac{\partial \Omega_{p_{i+1},N_{i+1}}}{\partial x}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i) & \geq & 0, \\ \frac{\partial \Omega_{p_{i+1},N_{i+1}}}{\partial y}(1-x_{i+1}, 1-\left\{\boldsymbol{F_{p,N}}(\boldsymbol{x})\right\}_i) & \leq & 0. \end{array} \tag{B.15}$$

If we also *assume* that

$$\begin{array}{rcll} b_{i,j}(\boldsymbol{x}) & \leq & 0, & \text{for all } j \in \{1,2,..,m\}, \\ f_{i,j}(\boldsymbol{x}) & \leq & 0, & \text{for all } j \in \{1,2,..,m\}, \end{array} \tag{B.16}$$

then we can conclude from (B.13) that both

$$\begin{array}{rcll} b_{i-1,j}(\boldsymbol{x}) & \leq & 0, & \text{for all } j \in \{1,2,..,m\}, \\ f_{i+1,j}(\boldsymbol{x}) & \leq & 0, & \text{for all } j \in \{1,2,..,m\}. \end{array} \tag{B.17}$$

This completes the demonstration of the statement in Expression (B.10).
Next we observe that:

$$\begin{array}{rcll} b_{m,j}(\boldsymbol{x}) & = & 0, & \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, j \in \{1,2,..,m\}, \\ f_{1,j}(\boldsymbol{x}) & = & 0, & \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, j \in \{1,2,..,m\}. \end{array} \tag{B.18}$$

A combination of the statements in Expressions (B.18) and (B.10) shows by means of induction that:

$$\begin{array}{rcll} b_{i,j}(\boldsymbol{x}) & \leq & 0, & \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, (i,j) \in \{1,2,..,m\}^2, \\ f_{i,j}(\boldsymbol{x}) & \leq & 0, & \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, (i,j) \in \{1,2,..,m\}^2. \end{array} \tag{B.19}$$

From the chain rule of differentiation of vector functions follows:

$$\boldsymbol{C}'_{\boldsymbol{p,N}}(\boldsymbol{x}) = (\boldsymbol{F_{p,N}} \circ \boldsymbol{B_{p,N}})'(\boldsymbol{x}) = \boldsymbol{F}'_{\boldsymbol{p,N}}(\boldsymbol{y}).\boldsymbol{B}'_{\boldsymbol{p,N}}(\boldsymbol{x}), \tag{B.20}$$

where:

$$\boldsymbol{y} = \boldsymbol{B_{p,N}}(\boldsymbol{x}). \tag{B.21}$$

Therefore:

$$c_{i,j}(\boldsymbol{x}) = \sum_{k=1}^{m} f_{i,k}(\boldsymbol{y}).b_{k,j}(\boldsymbol{x}). \tag{B.22}$$

Now we combine (B.22) with (B.19) and we get:

$$c_{i,j}(\boldsymbol{x}) \geq 0, \quad \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, (i,j) \in \{1,2,..,m\}^2. \tag{B.23}$$

This completes the proof of the lemma.

■

Next we define:

Definition 11

$$\boldsymbol{x} \leq \boldsymbol{y} \quad \overset{\text{def}}{\Leftrightarrow} \quad x_i \leq y_i, \quad \forall\, i \in \{1,2,..,m\}. \tag{B.24}$$

B.3 Proof of Monotonous Behaviour

Theorem 27

$$\begin{array}{rcll} \boldsymbol{p}^f(s+1) & \geq & \boldsymbol{p}^f(s), & \forall\, s \in I\!N, \\ \boldsymbol{p}^b(s+1) & \leq & \boldsymbol{p}^b(s), & \forall\, s \in I\!N. \end{array} \tag{B.25}$$

Proof. From Lemma 3:

$$c_{i,j}(\boldsymbol{x}) \geq 0, \quad \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, (i,j) \in \{1,2,..,m\}^2, \tag{B.26}$$

follows directly:

$$\boldsymbol{x} \leq \boldsymbol{y} \quad \Rightarrow \quad \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{x}) \leq \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{y}). \tag{B.27}$$

Therefore:

$$\begin{array}{rcl} \boldsymbol{p}^f(s-1) \leq \boldsymbol{p}^f(s) & \Rightarrow & \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{p}^f(s-1)) \leq \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{p}^f(s)) \quad \Rightarrow \\ & \Rightarrow & \boldsymbol{p}^f(s) \leq \boldsymbol{p}^f(s+1). \end{array} \tag{B.28}$$

From the initial condition follows:

$$\boldsymbol{p}^f(0) = \boldsymbol{0} \leq \boldsymbol{p}^f(1). \tag{B.29}$$

From (B.29) and (B.28) follows by means of induction:

$$\boldsymbol{p}^f(s+1) \geq \boldsymbol{p}^f(s), \quad \forall\, s \in I\!N. \tag{B.30}$$

From Lemma 3:

$$b_{i,j}(\boldsymbol{x}) \leq 0, \quad \forall\, \boldsymbol{x} \in (0,1)^m,\ \forall\, (i,j) \in \{1,2,..,m\}^2, \tag{B.31}$$

follows directly:

$$\boldsymbol{x} \leq \boldsymbol{y} \quad \Rightarrow \quad \boldsymbol{B}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{x}) \geq \boldsymbol{B}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{y}). \tag{B.32}$$

Therefore:

$$\begin{array}{rcl} \boldsymbol{p}^f(s) \leq \boldsymbol{p}^f(s+1) & \Rightarrow & \boldsymbol{B}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{p}^f(s)) \geq \boldsymbol{B}_{\boldsymbol{p},\boldsymbol{N}}(\boldsymbol{p}^f(s+1)) \quad \Rightarrow \\ & \Rightarrow & \boldsymbol{p}^b(s) \geq \boldsymbol{p}^b(s+1). \end{array} \tag{B.33}$$

From (B.30) and (B.33) follows:

$$\boldsymbol{p}^b(s+1) \leq \boldsymbol{p}^b(s), \quad \forall\, s \in I\!N. \tag{B.34}$$

This completes the proof of the theorem.

■

B.4 Proof of Existence

Theorem 28

$$\begin{array}{c} \boldsymbol{p}^f \stackrel{\text{def}}{=} \lim_{s\to\infty} \boldsymbol{p}^f(s), \\ \boldsymbol{p}^b \stackrel{\text{def}}{=} \lim_{s\to\infty} \boldsymbol{p}^b(s), \\ \text{both exist.} \end{array} \tag{B.35}$$

Proof. It is obvious that:

$$\begin{array}{l} \boldsymbol{p}^f(s) \leq \mathbf{1}, \quad \forall\ s\ \in\ I\!N, \\ \boldsymbol{p}^b(s) \geq \mathbf{0}, \quad \forall\ s\ \in\ I\!N. \end{array} \tag{B.36}$$

From Theorem 27 follows:

$$\begin{array}{l} \boldsymbol{p}^f(s+1) \geq \boldsymbol{p}^f(s), \quad \forall\ s\ \in\ I\!N, \\ \boldsymbol{p}^b(s+1) \leq \boldsymbol{p}^b(s), \quad \forall\ s\ \in\ I\!N. \end{array} \tag{B.37}$$

Therefore the forward sequence is 'increasing' and has an upper bound. The backward sequence is 'decreasing' and has a lower bound. We can conclude that both sequences converge in the maximum norm in $I\!R^m$ for instance. Because all norms in $I\!R^m$ are equivalent, both sequences will converge in any norm. This completes the proof.

■

B.5 Proof of Conservation of Flow

Theorem 29 *The combination of limit points $\boldsymbol{p}^f$ and $\boldsymbol{p}^b$ together satisfy all conditions of Expression (2.50), and therefore automatically satisfy Expression (2.45).*

Proof. It is obvious that these limit points satisfy (2.50), since they are constructed by means of this expression. That the limit points satisfy Expression (2.45) is to be proven.

First we define r_i:

$$r_i \stackrel{\text{def}}{=} R_{N_i}(p_{i-1}^f, p_i^b), \quad \forall\ i\ \in\ \{2,..,m\}. \tag{B.38}$$

Because of a basic property of function $R_N(x,y)$ defined in (2.20), it is automatically true that:

$$r_i = R_{N_i}(p_i^b, p_{i-1}^f), \quad \forall\ i\ \in\ \{2,..,m\}. \tag{B.39}$$

From (2.20), (2.48), (2.49) and (2.50) follows that:

$$\begin{array}{c} r_1 = p_1^b, \quad r_m = p_m^f, \\ p_i^f = \Omega_{p_i,N_i}(q_i^b, q_{i-1}^f) = \dfrac{p_i}{p_i^b} \cdot R_{N_i}(p_i^b, p_{i-1}^f) = \dfrac{p_i}{p_i^b} \cdot r_{i-1}, \\ p_i^b = \Omega_{p_i,N_{i+1}}(q_i^f, q_{i+1}^b) = \dfrac{p_i}{p_i^f} \cdot R_{N_{i+1}}(p_i^f, p_{i+1}^b) = \dfrac{p_i}{p_i^f} \cdot r_i. \end{array} \tag{B.40}$$

And thus:

$$\begin{aligned} r_{i-1} &= \frac{p_i^f \cdot p_i^b}{p_i}, \\ r_i &= \frac{p_i^f \cdot p_i^b}{p_i}. \end{aligned} \tag{B.41}$$

And therefore:

$$r_i = r_j, \quad \forall\, i,j \;\in\; \{2,3,..,m\}. \tag{B.42}$$

This completes the proof of the theorem.

■

B.6 Proof of Uniqueness

Theorem 30 *The solution of Equations (2.50) is unique.*

Proof. For our proof we first define a new function $Q_N(x,y)$ that has the following property:

$$R_N(x,y) = z \quad \Leftrightarrow \quad Q_N(x,z) = y \quad \Leftrightarrow \quad Q_N(y,z) = x \tag{B.43}$$

In order to show that such a function Q_N exists uniquely we have shown in Lemma 1 from Appendix A in Section A.4 that for two given constants $a \in (0,1]$ and $b \in [0,a]$ there exists a unique solution for $x \in [0,1]$ in the equation:

$$R_N(a,x) = b. \tag{B.44}$$

For the sake of convenient notation we will repeat the definition of the set V just as in definition 4 of the half unit square:

$$V \stackrel{\text{def}}{=} \{\ (x,y) \in [0,1] \times (0,1] \ \mid x \geq y\ \}. \tag{B.45}$$

As we have shown in Lemma 1, there exists a unique function Q_N that has the following properties:

$$\begin{aligned} Q_N \quad &: \quad V \;\mapsto\; [0,1]\,, \\ R_N(x, Q_N(x,y)) \;&\stackrel{\text{def}}{=}\; y, \qquad \forall\ (x,y) \in V. \end{aligned} \tag{B.46}$$

In the previous theorem we have proved that every solution of Equations (2.50) has the following property:

$$\begin{aligned} R \;&=\; r_i, \quad \forall i \in \{2,3,..,m\}, \\ r_i \;&=\; R_{N_i}(p_{i-1}^f, p_i^b), \quad \forall i \in \{2,3,..,m\}. \end{aligned} \tag{B.47}$$

From a basic property in Expression (2.42) follows:

$$R = \frac{p_i^f p_i^b}{p_i}, \quad \forall i \in \{2,3,..,m\}. \tag{B.48}$$

From Expressions (B.47) we can derive by means of the functions Q_N that:

$$p_i^b = Q_{N_i}(p_{i-1}^f, R), \tag{B.49}$$

and from Expression (B.48) we can derive that:

$$p_i^f = \frac{p_i R}{p_i^b}. \tag{B.50}$$

If we combine the last two expressions we get:

$$p_i^f = \frac{p_i R}{Q_{N_i}(p_{i-1}^f, R)}, \quad \forall i \in \{2, 3, .., m\}. \tag{B.51}$$

We also realize that for any solution of Equations (2.50) $p_1^f = p_1$ and $p_m^f = R$. This means that *all* solutions of Equations (2.50) are also solutions of the equations:

$$\begin{aligned} p_i^f &= \frac{p_i R}{Q_{N_i}(p_{i-1}^f, R)}, \quad \forall i \in \{2, 3, .., m\}. \\ p_1^f &= p_1, \\ p_m^f &= R. \end{aligned} \tag{B.52}$$

By means of the first two equations in Expression (B.52) we can derive all p_i^f recursively as a function of R.
First we define, for $a_1 = 1$, p_1^f as a function of R:

$$\begin{aligned} p_1^f &: (0, a_1] \mapsto [a_1 p_1, p_1], \\ p_1^f(R) &\stackrel{\text{def}}{=} p_1. \end{aligned} \tag{B.53}$$

For this function p_1^f it is obvious that

$$\begin{aligned} \lim_{R \downarrow 0} p_1^f(R) &= p_1 > 0, \\ p_1^f(a_1) &= a_1 p_1 \leq a_1, \\ \frac{dp_1^f}{dR}(R) &\leq 0, \quad \forall R \in (0, 1). \end{aligned}$$

Suppose we have defined p_{i-1}^f as a continuous function of R for some $a_{i-1} \in (0, 1]$ as follows:

$$p_{i-1}^f : (0, a_{i-1}] \mapsto [a_{i-1} p_{i-1}, p_{i-1}]. \tag{B.54}$$

And suppose that we know that

$$\begin{aligned} \lim_{R \downarrow 0} \; p_{i-1}^f(R) &= p_{i-1} > 0, \\ p_{i-1}^f(a_{i-1}) &= a_{i-1}p_{i-1} \leq a_{i-1}, \\ \frac{dp_{i-1}^f}{dR}(R) &\leq 0, \quad \forall R \in (0, a_{i-1}). \end{aligned} \tag{B.55}$$

Then it is easy to show by means of these properties of p_{i-1}^f, because of continuity, monotonicity and because the endpoints of p_{i-1}^f lie on either side of the function R, that there exists a *unique* solution for R in the equation

$$p_{i-1}^f(R) = R.$$

This unique solution of $R \in (0, 1]$ is then denoted by a_i:

$$p_{i-1}^f(a_i) = a_i. \tag{B.56}$$

Since a_i is in the domain of function p_{i-1}^f, which is $(0, a_{i-1}]$, it is obvious that

$$a_i \leq a_{i-1}.$$

It is easy to see by the properties of p_{i-1}^f that then for all $R \in (0, a_i]$ the following holds:

$$p_{i-1}^f(R) \geq R,$$

and therefore we conclude that for all $R \in (0, a_i]$ is $(p_{i-1}^f(R), R) \in V$.
This means that the following definition of the succeeding function $p_i^f(R)$ is well-defined:

$$\begin{aligned} p_i^f &: (0, a_i] \mapsto [a_i p_i, p_i], \\ p_i^f(R) &\stackrel{\text{def}}{=} \frac{p_i \, R}{Q_N(p_{i-1}^f(R), R)}. \end{aligned} \tag{B.57}$$

Again we see by "Q_N property 6" in Appendix A that

$$\begin{aligned} \lim_{R \downarrow 0} \; p_i^f(R) &= \lim_{R \downarrow 0} \; p_i \cdot \frac{R}{Q_N(p_{i-1}, R)}, \\ &= p_i \cdot 1, \\ &= p_i. \end{aligned}$$

We also see that

$$p_i^f(a_i) = \frac{p_i\ a_i}{Q_N(p_{i-1}^f(a_i), a_i)},$$

which reduces by means of Expression (B.56) to

$$p_i^f(a_i) = \frac{p_i\ a_i}{Q_N(a_i, a_i)}.$$

Because of "Q_N property 1" in Appendix A we conclude that $Q_N(a_i, a_i) = 1$, and thus

$$p_i^f(a_i) = p_i\ a_i.$$

Last but not least we derive that

$$\begin{aligned}\frac{dp_i^f}{dR}(R) &= \frac{d}{dR}\left(\frac{p_i\ R}{Q_N(p_{i-1}^f(R), R)}\right), \\ &= p_i \cdot \frac{\left[Q_N(p_{i-1}^f(R), R) - R.\frac{\partial Q_N}{\partial y}(p_{i-1}^f(R), R)\right] - R.\frac{\partial Q_N}{\partial x}(p_{i-1}^f(R), R) \cdot \frac{dp_{i-1}^f}{dR}(R)}{Q_N^2(p_{i-1}^f(R), R)}.\end{aligned} \tag{B.58}$$

Again we use the properties of Q_N in Appendix A.
Let $R \in (0, a_i)$. This means that $(p_i^f(R), R) \in U$ and therefore we can apply "Q_N property 5", substitute the pair $(p_i^f(R), R)$ for (x, y) and get

$$\frac{\partial Q_N}{\partial y}(p_i^f(R), R) > \frac{Q_N(p_i^f(R), R)}{R}.$$

This implies that

$$Q_N(p_{i-1}^f(R), R) - R.\frac{\partial Q_N}{\partial y}(p_{i-1}^f(R), R) < 0. \tag{B.59}$$

From "Q_N property 3" follows if we substitute the pair $(p_i^f(R), R)$ for (x, y) the following

$$\frac{\partial Q_N}{\partial x}(p_{i-1}^f(R), R) < 0,$$

and from the predecessor function p_{i-1}^f we know by assumption that

$$\frac{dp_{i-1}^f}{dR}(R) \leq 0, \quad \forall R \in (0, a_{i-1}).$$

These last two inequalities lead to the following inequality

$$-R.\frac{\partial Q_N}{\partial x}(p_{i-1}^f(R),R)\cdot\frac{dp_{i-1}^f}{dR}(R) \quad \leq \quad 0. \tag{B.60}$$

Now we can combine the inequalities of (B.59) and (B.60) and get for the numerator on the right side of Expression (B.58)

$$\left[Q_N(p_{i-1}^f(R),R)-R.\frac{\partial Q_N}{\partial y}(p_{i-1}^f(R),R)\right]-R.\frac{\partial Q_N}{\partial x}(p_{i-1}^f(R),R)\cdot\frac{dp_{i-1}^f}{dR}(R) \quad < \quad 0.$$

The denominator on the right side of Expression (B.58) obviously is positive and therefore we conclude that

$$\frac{dp_i^f}{dR}(R) \quad < \quad 0, \quad \forall R\in(0,a_i).$$

To resume our story, we started with the assumption that we have defined a function p_{i-1}^f for which has the properties expressed in (B.55) and we have shown that then by the definition in (B.57) we have a well-defined function p_i^f that has the same properties expressed in (B.55) but then for $i+1$ substituted for i.
By the induction argument and the initial condition in Expression (B.53) we conclude that we have found *unique* real numbers a_i, $i\in\{1,2,...,m\}$,

$$0 \;<\; a_m \;\leq\; a_{m-1} \;\leq\; \ldots \;\leq\; a_2 \;\leq\; a_1 \;=\; 1,$$

and we have defined functions

$$p_i^f \quad : \quad (0,a_i] \;\mapsto\; [a_ip_i,p_i], \qquad i\in\{1,2,...,m\},$$

for which the following properties hold

$$\lim_{R\downarrow 0}\; p_i^f(R) \;=\; p_i \;\geq\; 0,$$

$$p_i^f(a_i) \;=\; a_ip_i \;\leq\; a_i,$$

$$\frac{dp_i^f}{dR}(R) \;\leq\; 0, \quad \forall R\in(0,a_i).$$

Because of these properties all of the following equations for $i\in\{1,2,...,m\}$ have a *unique* solution for $R\in(0,1]$:

$$p_i^f(R) \quad = \quad R.$$

The solution for this equation is *unique* for $i=m$ in particular.
This implies that there exists exactly one *unique* $R\in(0,1]$ that satisfies the last equation

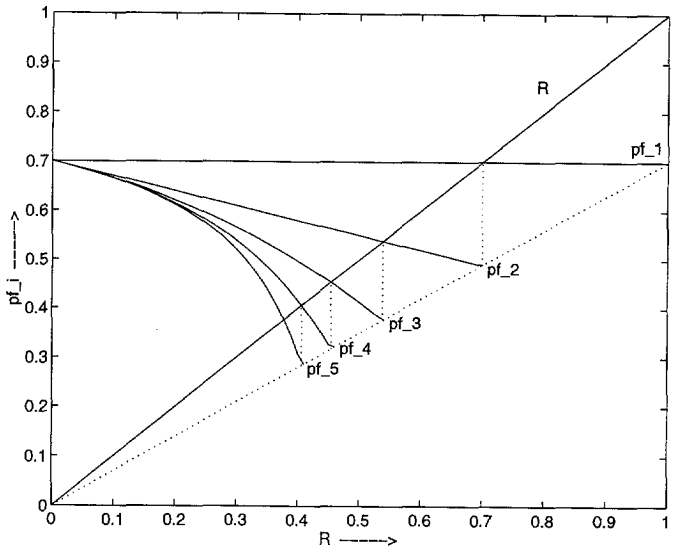

Figure B.1: A numerical example of functions $p_i^f(R)$ in case $p_i = 0.7$ and $N_i = 1$ for all i.

in Expression (B.52).
This completes the proof of the theorem.

■

Example.

For the sake of illustration we have numerically computed some functions $p_i^f(R)$ for a five machine production line that has $p_i = 0.7$ for all $i \in \{1, 2, 3, 4, 5\}$ and $N_i = 1$ for all $i \in \{2, 3, 4, 5\}$. The numerical result is shown in Figure B.1. We have taken an example in which all completion probabilities are equal. Therefore we can see that all the functions p_i^f for $i \in \{1, 2, 3, 4, 5\}$ start at the same point $(0, 0.7)$ and they all end on the same line (x, y), $y = 0.7\ x$. The numerical values for a_i, $i \in \{1, 2, 3, 4, 5\}$ are 1, 0.7000, 0.5385, 0.4541 and 0.4065 respectively. The final equation $p_5^f(R) = R$ has the numerical solution $R = 0.3774$, which means that the estimated overall average production rate R for this example is 0.3774 products per time slot. For $R = 0.3774$ we compute that p_i^f for $i \in \{1, 2, 3, 4, 5\}$ is 0.7000, 0.5868, 0.5140, 0.4502 and 0.3774 respectively.
We can see clearly in this example that the functions $p_i^f(R)$ are decreasing for all $i \in \{1, 2, 3, 4, 5\}$ and that begin and end points of $p_i^f(R)$ lie on either side of the line (x, y), $y = x$, for all $i \in \{1, 2, 3, 4, 5\}$.

□

Appendix C

Tables for the Car Lamp Production Line in Chapter 5

In Chapter 5 we have omitted tables because they would decrease the readibility of the text. Therefore we have decided to put these tables here in an extra appendix. This appendix is divided in two sections: a section that contains the tables that correspond to the text in Section 5.3, and a section that contains the tables that correspond to the text in Section 5.9.

C.1 Tables with the State sets of the Various Models and Sub-Models

In Chapter 5, in Section 5.3, we described several state sets of homogeneous discrete-time finite Markov chains. Each time that we defined a new state set (denoted by a symbol that contains a calligrafic $\mathcal{S}$) as a cartesian product of several sets, for the convenience of a shorter notation, we defined also a corresponding state set (denoted by a symbol that contains a calligraphic $\mathcal{R}$) that consists of (combinations of) natural numbers only. The correspondence between each pair of sets are described in the following tables, Table C.1 up to Table C.10.

State in $\mathcal{R}^{B_{2,1}}$	Corresponding state in $\mathcal{S}^{B_{2,1}}$					
	$B_{1,1}$	$M_{1,1}$	$B_{2,1}$	$B_{2,2}$	B_3	M_2
(c,1)	¬starved	up	c	¬starved	¬blocked	up
(c,2)	¬starved	up	c	¬starved	¬blocked	down
(c,3)	¬starved	up	c	¬starved	blocked	up
(c,4)	¬starved	up	c	¬starved	blocked	down
(c,5)	¬starved	up	c	starved	¬blocked	up
(c,6)	¬starved	up	c	starved	¬blocked	down
(c,7)	¬starved	up	c	starved	blocked	up
(c,8)	¬starved	up	c	starved	blocked	down
(c,9)	starved	up	c	¬starved	¬blocked	up
(c,10)	starved	up	c	¬starved	¬blocked	down
(c,11)	starved	up	c	¬starved	blocked	up
(c,12)	starved	up	c	¬starved	blocked	down
(c,13)	starved	up	c	starved	¬blocked	up
(c,14)	starved	up	c	starved	¬blocked	down
(c,15)	starved	up	c	starved	blocked	up
(c,16)	starved	up	c	starved	blocked	down
(c,17)	¬starved	down	c	¬starved	¬blocked	up
(c,18)	¬starved	down	c	¬starved	¬blocked	down
(c,19)	¬starved	down	c	¬starved	blocked	up
(c,20)	¬starved	down	c	¬starved	blocked	down
(c,21)	¬starved	down	c	starved	¬blocked	up
(c,22)	¬starved	down	c	starved	¬blocked	down
(c,23)	¬starved	down	c	starved	blocked	up
(c,24)	¬starved	down	c	starved	blocked	down
(c,25)	starved	down	c	¬starved	¬blocked	up
(c,26)	starved	down	c	¬starved	¬blocked	down
(c,27)	starved	down	c	¬starved	blocked	up
(c,28)	starved	down	c	¬starved	blocked	down
(c,29)	starved	down	c	starved	¬blocked	up
(c,30)	starved	down	c	starved	¬blocked	down
(c,31)	starved	down	c	starved	blocked	up
(c,32)	starved	down	c	starved	blocked	down

Table C.1: The one to one correspondence of $\mathcal{S}^{B_{2,1}}$ and $\mathcal{R}^{B_{2,1}}$.

State in $\mathcal{R}^{B_{2,1},\text{in}}$	Corresponding state in $\mathcal{S}^{B_{2,1},\text{in}}$	
	$B_{1,1}$	$M_{1,1}$
1	¬starved	up
2	¬starved	down
3	starved	up
4	starved	down

Table C.2: Corresponding states of the input sub-model for buffer $B_{2,1}$.

State in $\mathcal{R}^{B_{2,1},\text{out}}$	Corresponding state in $\mathcal{S}^{B_{2,1},\text{out}}$		
	$B_{2,2}$	B_3	M_2
1	¬starved	¬blocked	up
2	¬starved	¬blocked	down
3	¬starved	blocked	up
4	¬starved	blocked	down
5	starved	¬blocked	up
6	starved	¬blocked	down
7	starved	blocked	up
8	starved	blocked	down

Table C.3: Corresponding states of the output sub-model for buffer $B_{2,1}$.

State in $\mathcal{R}^{B_{2,2}}$	Corresponding state in $\mathcal{S}^{B_{2,2}}$					
	$B_{1,2}$	$M_{1,2}$	$B_{2,2}$	$B_{2,1}$	B_3	M_2
(c,1)	¬starved	up	c	¬starved	¬blocked	up
(c,2)	¬starved	up	c	¬starved	¬blocked	down
(c,3)	¬starved	up	c	¬starved	blocked	up
(c,4)	¬starved	up	c	¬starved	blocked	down
(c,5)	¬starved	up	c	starved	¬blocked	up
(c,6)	¬starved	up	c	starved	¬blocked	down
(c,7)	¬starved	up	c	starved	blocked	up
(c,8)	¬starved	up	c	starved	blocked	down
(c,9)	starved	up	c	¬starved	¬blocked	up
(c,10)	starved	up	c	¬starved	¬blocked	down
(c,11)	starved	up	c	¬starved	blocked	up
(c,12)	starved	up	c	¬starved	blocked	down
(c,13)	starved	up	c	starved	¬blocked	up
(c,14)	starved	up	c	starved	¬blocked	down
(c,15)	starved	up	c	starved	blocked	up
(c,16)	starved	up	c	starved	blocked	down
(c,17)	¬starved	down	c	¬starved	¬blocked	up
(c,18)	¬starved	down	c	¬starved	¬blocked	down
(c,19)	¬starved	down	c	¬starved	blocked	up
(c,20)	¬starved	down	c	¬starved	blocked	down
(c,21)	¬starved	down	c	starved	¬blocked	up
(c,22)	¬starved	down	c	starved	¬blocked	down
(c,23)	¬starved	down	c	starved	blocked	up
(c,24)	¬starved	down	c	starved	blocked	down
(c,25)	starved	down	c	¬starved	¬blocked	up
(c,26)	starved	down	c	¬starved	¬blocked	down
(c,27)	starved	down	c	¬starved	blocked	up
(c,28)	starved	down	c	¬starved	blocked	down
(c,29)	starved	down	c	starved	¬blocked	up
(c,30)	starved	down	c	starved	¬blocked	down
(c,31)	starved	down	c	starved	blocked	up
(c,32)	starved	down	c	starved	blocked	down

Table C.4: The one to one correspondence of $\mathcal{S}^{B_{2,2}}$ and $\mathcal{R}^{B_{2,2}}$.

State in $\mathcal{R}^{B_{2,2},\text{in}}$	Corresponding state in $\mathcal{S}^{B_{2,2},\text{in}}$	
	$B_{1,2}$	$M_{1,2}$
1	¬starved	up
2	¬starved	down
3	starved	up
4	starved	down

Table C.5: Corresponding states of the input sub-model for buffer $B_{2,2}$.

State in $\mathcal{R}^{B_{2,2},\text{out}}$	Corresponding state in $\mathcal{S}^{B_{2,2},\text{out}}$		
	$B_{2,1}$	B_3	M_2
1	¬starved	¬blocked	up
2	¬starved	¬blocked	down
3	¬starved	blocked	up
4	¬starved	blocked	down
5	starved	¬blocked	up
6	starved	¬blocked	down
7	starved	blocked	up
8	starved	blocked	down

Table C.6: Corresponding states of the output sub-model for buffer $B_{2,2}$.

State in $\mathcal{R}^{B_3}$	Corresponding state in $\mathcal{S}^{B_3}$					
	$B_{2,1}$	$B_{2,2}$	M_2	B_3	B_4	M_3
(c,1)	¬starved	¬starved	up	c	¬blocked	up
(c,2)	¬starved	¬starved	up	c	¬blocked	down
(c,3)	¬starved	¬starved	up	c	blocked	up
(c,4)	¬starved	¬starved	up	c	blocked	down
(c,5)	¬starved	¬starved	down	c	¬blocked	up
(c,6)	¬starved	¬starved	down	c	¬blocked	down
(c,7)	¬starved	¬starved	down	c	blocked	up
(c,8)	¬starved	¬starved	down	c	blocked	down
(c,9)	¬starved	starved	up	c	¬blocked	up
(c,10)	¬starved	starved	up	c	¬blocked	down
(c,11)	¬starved	starved	up	c	blocked	up
(c,12)	¬starved	starved	up	c	blocked	down
(c,13)	¬starved	starved	down	c	¬blocked	up
(c,14)	¬starved	starved	down	c	¬blocked	down
(c,15)	¬starved	starved	down	c	blocked	up
(c,16)	¬starved	starved	down	c	blocked	down
(c,17)	starved	¬starved	up	c	¬blocked	up
(c,18)	starved	¬starved	up	c	¬blocked	down
(c,19)	starved	¬starved	up	c	blocked	up
(c,20)	starved	¬starved	up	c	blocked	down
(c,21)	starved	¬starved	down	c	¬blocked	up
(c,22)	starved	¬starved	down	c	¬blocked	down
(c,23)	starved	¬starved	down	c	blocked	up
(c,24)	starved	¬starved	down	c	blocked	down
(c,25)	starved	starved	up	c	¬blocked	up
(c,26)	starved	starved	up	c	¬blocked	down
(c,27)	starved	starved	up	c	blocked	up
(c,28)	starved	starved	up	c	blocked	down
(c,29)	starved	starved	down	c	¬blocked	up
(c,30)	starved	starved	down	c	¬blocked	down
(c,31)	starved	starved	down	c	blocked	up
(c,32)	starved	starved	down	c	blocked	down

Table C.7: The one to one correspondence of $\mathcal{S}^{B_3}$ and $\mathcal{R}^{B_3}$.

State in $\mathcal{R}^{B_3,\text{in}}$	Corresponding state in $\mathcal{S}^{B_3,\text{in}}$		
	$B_{2,1}$	$B_{2,2}$	M_2
1	¬starved	¬starved	up
2	¬starved	¬starved	down
3	¬starved	starved	up
4	¬starved	starved	down
5	starved	¬starved	up
6	starved	¬starved	down
7	starved	starved	up
8	starved	starved	down

Table C.8: Corresponding states of the input sub-model for buffer B_3.

State in $\mathcal{R}^{B_3,\text{out}}$	Corresponding state in $\mathcal{S}^{B_3,\text{out}}$	
	B_4	M_3
1	¬blocked	up
2	¬blocked	down
3	blocked	up
4	blocked	down

Table C.9: Corresponding states of the output sub-model for buffer B_3.

State in $\mathcal{R}^{\text{overlap}}$	Corresponding state in $\mathcal{S}^{\text{overlap}}$			
	$B_{2,1}$	$B_{2,2}$	B_3	M_2
1	¬starved	¬starved	¬blocked	up
2	¬starved	¬starved	¬blocked	down
3	¬starved	¬starved	blocked	up
4	¬starved	¬starved	blocked	down
5	¬starved	starved	¬blocked	up
6	¬starved	starved	¬blocked	down
7	¬starved	starved	blocked	up
8	¬starved	starved	blocked	down
9	starved	¬starved	¬blocked	up
10	starved	¬starved	¬blocked	down
11	starved	¬starved	blocked	up
12	starved	¬starved	blocked	down
13	starved	starved	¬blocked	up
14	starved	starved	¬blocked	down
15	starved	starved	blocked	up
16	starved	starved	blocked	down

Table C.10: Corresponding states of the model for the overlap.

C.2 Tables with the Results from the Algorithm for the Car Lamp Production Line

After 12 iteration cycles with the algorithm described in Section 5.8 for the car lamp production line with the parameters as in Figure 5.24 we obtain the results shown in the following tables, Table C.11 up to Table C.15.

	Value
$y_{2,1}(12)$	0.736260
$x_3(12)$	0.680686
$y_{2,2}(12)$	0.733072
$\bar{x}_3(12)$	0.680686

Table C.11: The average input and output rates as a result of the algorithm.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	9954	0167	0418	0007	0258	0004	0011	0000	0081	0001	0003	0000	0002	0000	0000	0000
2	0028	9833	0001	0413	0001	0254	0000	0011	0000	0080	0000	0003	0000	0002	0000	0000
3	0003	0000	9540	0160	0000	0000	0247	0004	0000	0000	0078	0001	0000	0000	0002	0000
4	0000	0000	0027	9420	0000	0000	0001	0244	0000	0000	0000	0077	0000	0000	0000	0002
5	0013	0000	0001	0000	9714	0163	0408	0007	0000	0000	0000	0000	0079	0001	0003	0000
6	0000	0000	0000	0000	0028	9578	0001	0402	0000	0000	0000	0000	0000	0078	0000	0003
7	0000	0000	0013	0000	0000	0000	9306	0156	0000	0000	0000	0000	0000	0000	0076	0001
8	0000	0000	0000	0000	0000	0000	0027	9176	0000	0000	0000	0000	0000	0000	0000	0075
9	0002	0000	0000	0000	0000	0000	0000	0000	9874	0166	0415	0007	0256	0004	0011	0000
10	0000	0000	0000	0000	0000	0000	0000	0000	0028	9752	0001	0410	0001	0252	0000	0011
11	0000	0000	0000	0000	0000	0000	0000	0000	0003	0000	9462	0159	0000	0000	0245	0004
12	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9343	0000	0000	0001	0242
13	0000	0000	0000	0000	0000	0000	0000	0000	0013	0000	0001	0000	9634	0162	0405	0007
14	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9500	0001	0399
15	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0013	0000	0000	0000	9230	0155
16	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0026	9101
	7968	1408	0056	0003	0384	0026	0002	0000	0132	0015	0001	0000	0005	0000	0000	0000

Table C.12: The Markov transition matrix $P_{2,1}^{\text{overlap}}(12)$ and the stationary distribution.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	9954	0167	0418	0007	0258	0004	0011	0000	0081	0001	0003	0000	0002	0000	0000	0000
2	0028	9833	0001	0413	0001	0254	0000	0011	0000	0080	0000	0003	0000	0002	0000	0000
3	0003	0000	9539	0160	0000	0000	0247	0004	0000	0000	0078	0001	0000	0000	0002	0000
4	0000	0000	0027	9420	0000	0000	0001	0244	0000	0000	0000	0077	0000	0000	0000	0002
5	0013	0000	0001	0000	9714	0163	0408	0007	0000	0000	0000	0000	0079	0001	0003	0000
6	0000	0000	0000	0000	0028	9578	0001	0402	0000	0000	0000	0000	0000	0078	0000	0003
7	0000	0000	0013	0000	0000	0000	9306	0156	0000	0000	0000	0000	0000	0000	0076	0001
8	0000	0000	0000	0000	0000	0000	0027	9176	0000	0000	0000	0000	0000	0000	0000	0075
9	0002	0000	0000	0000	0000	0000	0000	0000	9877	0166	0415	0007	0256	0004	0011	0000
10	0000	0000	0000	0000	0000	0000	0000	0000	0028	9752	0001	0410	0001	0252	0000	0011
11	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9462	0159	0000	0000	0245	0004
12	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9343	0000	0000	0001	0242
13	0000	0000	0000	0000	0000	0000	0000	0000	0013	0000	0001	0000	9634	0162	0405	0007
14	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9500	0001	0399
15	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0013	0000	0000	0000	9230	0155
16	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0026	9101
	7986	1408	0057	0003	0385	0026	0001	0000	0132	0016	0000	0000	0005	0000	0000	0000

Table C.13: The Markov transition matrix $P_3^{\text{overlap}}(12)$ and the stationary distribution.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	9953	0167	0419	0007	0258	0004	0011	0000	0082	0001	0003	0000	0002	0000	0000	0000
2	0028	9833	0001	0413	0001	0254	0000	0011	0000	0080	0000	0003	0000	0002	0000	0000
3	0003	0000	9551	0160	0000	0000	0247	0004	0000	0000	0078	0001	0000	0000	0002	0000
4	0000	0000	0027	9420	0000	0000	0001	0244	0000	0000	0000	0077	0000	0000	0000	0002
5	0014	0000	0000	0000	9709	0163	0408	0007	0000	0000	0000	0000	0079	0001	0003	0000
6	0000	0000	0000	0000	0028	9578	0001	0402	0000	0000	0000	0000	0000	0078	0000	0003
7	0000	0000	0000	0000	0003	0000	9304	0156	0000	0000	0000	0000	0000	0000	0076	0001
8	0000	0000	0000	0000	0000	0000	0027	9176	0000	0000	0000	0000	0000	0000	0000	0075
9	0001	0000	0000	0000	0000	0000	0000	0000	9890	0166	0415	0007	0256	0004	0011	0000
10	0000	0000	0000	0000	0000	0000	0000	0000	0028	9752	0001	0410	0001	0252	0000	0011
11	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9475	0159	0000	0000	0245	0004
12	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9343	0000	0000	0001	0242
13	0000	0000	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9634	0162	0405	0007
14	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9500	0001	0399
15	0000	0000	0000	0000	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9230	0155
16	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0026	9101
	7964	1408	0057	0003	0388	0026	0002	0000	0136	0016	0000	0000	0002	0000	0000	0000

Table C.14: The Markov transition matrix $P_{2,2}^{\text{overlap}}(12)$ and the stationary distribution.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	9953	0167	0418	0007	0258	0004	0011	0000	0082	0001	0003	0000	0002	0000	0000	0000
2	0028	9833	0001	0413	0001	0254	0000	0011	0000	0080	0000	0003	0000	0002	0000	0000
3	0003	0000	9538	0160	0000	0000	0247	0004	0000	0000	0078	0001	0000	0000	0002	0000
4	0000	0000	0027	9420	0000	0000	0001	0244	0000	0000	0000	0077	0000	0000	0000	0002
5	0013	0000	0001	0000	9712	0163	0408	0007	0000	0000	0000	0000	0079	0001	0003	0000
6	0000	0000	0000	0000	0028	9578	0001	0402	0000	0000	0000	0000	0000	0078	0000	0003
7	0000	0000	0013	0000	0000	0000	9304	0156	0000	0000	0000	0000	0000	0000	0076	0001
8	0000	0000	0000	0000	0000	0000	0027	9176	0000	0000	0000	0000	0000	0000	0000	0075
9	0001	0000	0000	0000	0000	0000	0000	0000	9890	0166	0415	0007	0256	0004	0011	0000
10	0000	0000	0000	0000	0000	0000	0000	0000	0028	9752	0001	0410	0001	0252	0000	0011
11	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9475	0159	0000	0000	0245	0004
12	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9343	0000	0000	0001	0242
13	0000	0000	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9634	0162	0405	0007
14	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0027	9500	0001	0399
15	0000	0000	0000	0000	0000	0000	0001	0000	0000	0000	0000	0000	0000	0000	9230	0155
16	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0026	9101
	7964	1408	0056	0003	0388	0026	0001	0000	0136	0016	0000	0000	0002	0000	0000	0000

Table C.15: The Markov transition matrix $\bar{P}_3^{\text{overlap}}(12)$ and the stationary distribution.

Bibliography

[All90] Arnold O. Allen. *Probability, Statistics, and Queueing Theory, With Computer Science Applications.* Academic Press, Inc., second edition, 1990.

[BCOQ92] F. Baccelli, G. Cohen, G.J. Olsder, and J-P. Quadrat. *Synchronization and Linearity, An Algebra for Discrete Event Systems.* John Wiley & Sons Ltd., 1992.

[Blo89] Gunnar Blom. *Probability and Statistics, Theory and Applications.* Springer-Verlag New York Inc., 1989.

[BOW91] Erik van Bracht, Geert Jan Olsder, and Jacob van der Woude. Analysis and synthesis of serial production lines. In *Proceedings of the 30th Conference on Decision and Control*, volume 3, pages 2168–2173, December 1991.

[Bra91] Erik van Bracht. Analysis and synthesis of serial production lines. Master's thesis, Delft University of Technology, Faculty of Aplied Mathematics and Computer Science, 1991.

[Bra93] Erik van Bracht. Performance analysis of a serial production line with machine breakdowns. Technical Report 93-123, Delft University of Technology, Faculty of Technical Mathematics and Informatics, Delft, 1993.

[Bra94] Erik van Bracht. Performance analysis of a branched production line with machine breakdowns and product failures. Technical Report 94-54, Delft University of Technology, Faculty of Technical Mathematics and Informatics, Delft, 1994.

[Bra95] Erik van Bracht. Performance analysis of a serial production line with machine breakdowns. In *Proceedings of the 1995 INRIA / IEEE Symposium on Emerging Technologies and Factory Automation*, volume 3, pages 417–424, October 1995.

[Cas93] Christos G. Cassandras. *Discrete Event Systems; Modeling and Performance Analysis.* Irwin, Homewood, 1993.

[Coh69] J.W. Cohen. *The Single Server Queue.* North-Holland Publishing Company, Amsterdam-London, 1969.

[DF93] Yves Dallery and Yannick Frein. On decomposition methods for tandem queueing networks with blocking. *Operations Research*, 41(2):386–399, 1993.

[DG92] Yves Dallery and Stanley B. Gershwin. Manufacturing flow line systems: A review of models and analytical results. *Queueing Systems*, (12):3–94, 1992.

[Dij89] Nico M. van Dijk. A simple bounding methodology for non-product-form queueing networks. *Queueing Networks with Blocking, Proceedings of the First International Workshop held in Raleigh, North Carolina, May 20-21, 1988*, pages 3–17, 1989.

[Dij93] Nico M. van Dijk. *Queueing Networks and Product Forms; A Systems Approach.* Wiley, Chichester, 1993.

[DRXL88] Y. Dallery, David R., and Xie X.-L. An efficient algorithm for analysis of transfer lines with unreliable machines and finite buffers. *IIE Transactions*, (20):280–283, 1988.

[Ger89] Stanley B. Gershwin. An efficient decomposition algorithm for unreliable tandem queuing systems with finite buffers. *Queueing Networks with Blocking, Proceedings of the First International Workshop held in Raleigh, North Carolina, May 20-21, 1988*, pages 127–146, 1989.

[Ger93] Stanley B. Gershwin. *Manufacturing Systems Engineering.* PTR Prentice-Hall, Englewood Cliffs, 1993.

[Goo88] Roe Goodman. *Introduction to Stochastic Models.* The Benjamin/Cummings Publishing Company, Inc., 1988.

[GS87] G.R. Grimmett and D.R. Stirzaker. *Probability and Random Processes.* Oxford University Press, 1987.

[HB67] Frederick S. Hillier and Ronald W. Boling. Finite queues in series with exponential or erlang service times - a numerical approach. *Operations Research*, (16):286–303, 1967.

[HC91] Yu-Chi Ho and Xi-Ren Cao. *Perturbation Analysis of Discrete Event Dynamic Systems.* Kluwer, Boston, 1991.

[HET79] Y.C. Ho, M.A. Eyler, and Chien T.T. A gradient technique for general buffer storage design in a production line. *Int. J. Prod. Res.*, 17(6):557–580, 1979.

[HL64] J.L. Hodges and E.L. Lehmann. *Basic Concepts of Probability and Statistics.* Holden-Day Inc., San Fransisco, 1964.

[MT90] S.M. Meerkov and F. Top. Asymptotically reliable serial production lines: Analysis, synthesis and a case study. In *Proceedings of the 11th IFAC World Congress*, pages 2–7, August 1990.

[Neu81] Marcel F. Neuts. *Matrix-Geometric Solutions in Stochastic Models, An Algorithmic Approach.* The Johns Hopkins University Press, 1981.

[PA89] H.G. Perros and T. Altiok. *Queueing Networks with Blocking, Proceedings of the First International Workshop held in Raleigh, North Carolina, May 20-21, 1988.* North Holland, 1989.

[Rao76] Nori Prakasa Rao. A generalization of the 'bowl phenomenon' in series production systems. *Int. J. Prod. Res.*, 14(4):437–443, 1976.

[Sha90] O.P. Sharma. *Markovian Queues.* Ellis Horwood Limited, 1990.

[Wal88] J. Walrand. *An Introduction to Queueing Networks.* Prentice Hall, Englewood Cliffs, N.J., 1988.

Samenvatting (Summary in Dutch)

Dit proefschrift is het resultaat van vijf jaar onderzoek aan de Technische Universiteit Delft. Het onderzoek is gestart in een academisch onderzoeksproject waarin verschillende Europese universiteiten deelnamen in samenwerking met een bekende producent van lampen. Het doel van dit onderzoeksproject was het optimaliseren van een bestaand autolampen produktieproces.

Bij nadere beschouwing blijkt dat het autolampen produktieproces wordt gerealiseerd door een produktielijn waarin metalen en glazen onderdelen in verschillende fasen worden bewerkt door een aantal machines. Voor iedere fase onderscheiden we een aparte machine. Tussen opeenvolgende bewerkingen door opeenvolgende machines worden de onderdelen in een wachtrij gezet, waarin ze wachten tot de volgende machine klaar is met het bewerken van alle voorgaande onderdelen. Het toelaten van wachtrijen is nuttig voor een zo groot mogelijke ontkoppeling van de verschillende bewerkingen in de produktielijn. Door deze ontkoppeling is het mogelijk dat een machine door een ongewenst probleem een beperkte tijd stil staat terwijl de rest van de machines hun werkzaamheden voortzetten.

Bij een wiskundige beschrijving van zo'n produktielijn stuiten we onherroepelijk op het feit dat de gebeurtenissen in de lijn afhangen van puur toeval. Er treden hoofdzakelijk de volgende soorten toeval op in deze produktielijn:

- Op ieder moment kunnen er in werkende machines ongewenste fouten optreden, waardoor de machine stil komt te staan totdat een "operator" komt om de fout te herstellen. Op deze wijze wordt zowel de aaneengesloten tijd dat een machine werkt als de aaneengesloten tijd dat een machine ongewenst stilstaat door toeval bepaald. Een simpel wiskundig model wordt gebruikt waarin alleen de gemiddelde aaneengesloten werkzame tijd en de gemiddelde aaneengesloten tijd van stilstand per machine als belangrijkste variabelen een rol spelen.

- In ieder onderdeel dat door een machine wordt bewerkt kunnen fouten optreden tijdens de bewerking. Vaak kan deze fout eenvoudig worden gesignaleerd. Als zo'n fout wordt gesignaleerd is het mogelijk het betreffende onderdeel uit het produktieproces te nemen. Deze fouten in de bewerking van produkten worden op zodanige manier gemodelleerd dat alleen de gemiddelde uitval over een lange periode per machine als variabele een rol speelt.

Doordat toeval de gebeurtenissen in de produktielijn bepalen moeten we voor een wiskundige beschrijving van de wachtrijen in de produktielijn gebruik maken van zogenaamde stochastische processen. Voor de wiskundige beschrijving van zulke stochastische processen van wachtende voorwerpen in een netwerk van wachtrijen komen we automatisch terecht in het gebied van de wachtrijnetwerken.

Wachtrijnetwerken bestaan in vele verschillende vormen. Als we de wachtrijen in de autolampen produktielijn nader bekijken constateren we dat hun capaciteit beperkt is. Dit heeft als gevolg dat wachtrijen tijdens het produktieproces vol kunnen raken waardoor voorafgaande machines stil komen te staan doordat ze de bewerkte produkten niet meer kwijt kunnen raken. Dit fenomeen wordt met de term "blokkering" aangeduid. De conclusie is dus dat we voor een wiskundige beschrijving ons moeten bezighouden met *blokkerende* wachtrijnetwerken.

In de autolampen produktielijn is er een machine die onderdelen uit twee verschillende wachtrijen tegelijk oppakt. Tijdens de bewerking worden deze twee onderdelen samengevoegd en als één onderdeel in een volgende wachtrij geplaatst. Deze procedure word aangeduid met de term "samenvoeging".

Hoewel er in de autolampen produktielijn sprake is van samenvoeging proberen we eerst inzicht te krijgen in meer simpele seriële produktielijnen. Het blijkt mogelijk om seriële produktielijnen op te delen in een verzameling simpele produktielijnen met ieder slechts twee machines. Door het bepalen van het stationaire gedrag van deze twee-machine produktielijntjes kunnen we zo conclusies trekken over het het stationaire gedrag van de hele produktielijn. Een aantal methoden van opdeling van seriële produktielijnen in kleine twee-machine produktielijntjes zijn in de laatste dertig jaar geïntroduceerd en mogen als klassiek worden beschouwd. Voor zover bekend gaat het hier echter altijd om machines met continu verdeelde bewerkingstijden. Omdat we in de autolampen produktielijn te maken hebben met machines die met een *vast interval* produkten afleveren, zogenaamde geïndexeerde machines, hebben we te hier niet te maken met continu maar met discreet verdeelde bewerkingstijden. Om deze discreet verdeelde bewerkingstijden beter te modelleren hebben we de klassieke methoden van opdeling in twee-machine produktielijntjes grotendeels gehandhaafd, maar ook met succes aangepast zodat zij werken met de discrete geometrische verdelingen. De nieuwe methode voor geometrisch verdeelde bewerkingstijden hebben we aan een zeer grondige wiskundige analyse onderworpen. Het is ons zodoende gelukt om existentie, eenduidigheid, convergentie en "behoud van gemiddelde doorstroming" te bewijzen voor de oplossing van deze nieuwe methode. (Behoud van gemiddelde doorstroming is een algemene eigenschap van seriële produktielijnen die zegt dat over een lange tijd iedere machine in de lijn gemiddeld evenveel onderdelen bewerkt.)

Voor de modellering van de autolampen produktielijn is deze modellering met behulp van geometrische verdelingen echter nog lang niet voldoende. Het is met de geometrisch

verdeelde bewerkingstijden namelijk niet mogelijk om de stilstand en reparatie van machines te beschrijven. Daarom proberen we allereerst inzicht te verkrijgen in seriële produktielijnen waarin *alleen* stilstand en reparatie van machines een rol spelen. Voor dit soort seriële produktielijnen hebben we een geheel nieuwe methode ontwikkeld die intensief gebruik maakt van een wiskundige methode om stochastische processen te reduceren door het indelen van de grote toestandsruimte in minder toestandsgroepen. Vele voorafgaande ideeën, zoals opdeling van seriële produktielijnen in kleine twee-machine produktielijntjes, zullen in deze nieuwe methode in een iets andere vorm worden hergebruikt. Deze nieuwe methode blijkt in voorbeeldsituaties zeer goed in staat te zijn de stilstand en reparatie van machines in een produktielijn te modelleren. De wiskundige "elegantie" van deze methode laat echter wel wat te wensen over. Hoewel het gebruikte concept van reductie van stochastische processen zeer eenvoudig is, is de complexiteit van de methode dermate hoog dat we er niet in zijn geslaagd existentie, eenduidigheid, convergentie en "behoud van gemiddelde doorstroming" te bewijzen. Wel is het duidelijk dat de resultaten van deze methode niet *precies* voldoen aan het "behoud van gemiddelde doorstroming". Dit is echter geen praktisch probleem omdat *alle bekende methoden* een *benadering* van het stationair gedrag bepalen. Zolang de gemiddelde doorstromingen voor de machines in de produktielijn elkaar dicht naderen kan er nog steeds sprake zijn van een goede benadering van het stationaire gedrag.

Vervolgens proberen we de modellen die zijn gebaseerd op discrete geometrische verdeelde bewerkingstijden en de modellen die zijn gebaseerd op stilstand en reparatie van machines te combineren en komen zo tot modellen van *seriële* produktielijnen die beide stochastische eigenschappen in zich verenigen. We ontwikkelen zodoende een methode waarmee we het stationaire gedrag kunnen bepalen van *seriële* produktielijnen waarvan de machines zowel een discrete geometrisch verdeelde bewerkingstijd hebben als aan stilstand en reparatie onderhevig zijn.

De laatste stap die we moeten zetten voor een wiskundig model van de autolampen produktielijn is de stap van een *seriële* produktielijn naar een *non-seriële* configuratie met samenvoeging. Deze stap word genomen door opnieuw een aanpassing en uitbreiding van de bestaande methoden voor seriële configuraties. Het resultaat is de mogelijkheid om een goede benadering te krijgen van het stationair gedrag van de autolampen produktielijn op basis van de volgende vier verschillende parameters per machine: de gemiddelde aaneengesloten werkende tijd, de gemiddelde aaneengesloten reparatietijd, gemiddelde uitval van onbruikbare onderdelen en de machinesnelheid. Naast deze parameters voor de machines is er nog één extra parameter per wachtrij: de maximale wachtrijlengte. Aan de hand van de waarden van deze parameters zijn we nu in staat het stationaire gedrag van de produktielijn te benaderen, inclusief de gemiddelde produktie van de gehele produktielijn. Door opnieuw berekenen na het aanbrengen van kleine variaties in de parameters kunnen we uitspraken doen over de gevoeligheid van de gemiddelde produktie voor deze variaties. Deze gevoeligheden kunnen worden gebruikt ter ondersteuning van besluiten die worden

genomen voor veranderingen in de autolampen produktielijn.

Naast deze lange weg waarlangs we een wiskundig model afleiden voor de berekening van het stationair gedrag van deze specifieke produktielijn, wordt ook meer fundamenteel onderzoek beschreven in een aanverwante richting die in principe betere beschrijvingen kan geven van wachtrijnetwerken. Deze richting is de richting van de stochastische systemen met discrete gebeurtenissen, in het Engels "stochastic timed event graphs" genoemd, die een subklasse zijn van de zogenaamde Petri netten. Het is mogelijk om een betere beschrijving van de autolampen produktielijn te maken met behulp van deze systemen. Echter, de berekening van het stationaire gedrag van zulke systemen zijn zeer gecompliceerd. In dit proefschrift wordt een voorbeeld behandeld van een zeer eenvoudig twee-dimensionaal autonoom "stochastic timed event graph". Het blijkt dat het zelfs daar nog onmogelijk is het stationaire gedrag van zo'n eenvoudig systeem in een expliciete vorm uit te drukken. Het zal daarom duidelijk zijn dat exacte berekening van het stationaire gedrag van een soortgelijk systeem voor de complexe autolampen produktielijn buiten het bereik is van de huidige wetenschap.

In dit proefschrift wordt toch een begin gemaakt met het modelleren van blokkerende seriële produktielijnen door middel van "stochastic timed event graphs". Het blijkt dat een zeer elegante beschrijving mogelijk is waarmee een zeer uitgebreide verzameling van seriële blokkerende wachtrijnetwerken met *willekeurige verdeelde service tijden* kunnen worden beschreven.

Ook deze modellen voor wachtrijnetwerken met behulp van "stochastic timed event graphs" worden vervolgens opgedeeld in kleinere systemen wij *Wait & Switch Systems* noemen. Door grondig onderzoek van deze Wait & Switch Systems blijkt het toch mogelijk het stationaire gedrag ervan te bepalen. Door gebruik te maken van de stationaire verdelingen van de diverse Wait & Switch Systems kunnen we zodoende een stationaire verdeling van de hele produktielijn benaderen. Voorbeelden van deze methode zijn nog niet behandeld in dit proefschrift. Wel is een algoritme beschreven waarmee deze stationaire verdeling met behulp van de Wait & Switch Systems kan worden benaderd. De auteur van dit proefschrift ziet het als een enorme uitdaging om een exacte stationaire verdeling van het model te bepalen. Het algoritme blijkt namelijk wel een goede benadering, maar zeker nog geen exacte stationaire verdeling van het stochastische model op te leveren. Verder onderzoek in deze richting lijkt zeer nuttig en kan misschien leiden tot een exacte oplossing van het model dat hier wordt beschreven. Dit zou een doorbraak kunnen betekenen in de beschrijving en berekening van het stationair gedrag van (blokkerende) wachtrijnetwerken.

Curriculum Vitae (in Dutch)

De schrijver van dit proefschrift werd op 13 juni 1968 om 8 uur 's ochtends te Gouda geboren, alwaar hij gedurende de volgende tijden de volgende opleidingen succesvol heeft doorlopen: twee jaar kleuterschool Sneeuwwitje, zes jaar basisschool De Springplank en zes jaar Voorbereidend Wetenschappelijk Onderwijs op het Rijksscholengemeenschap ter plaatse. Zodoende was hij met het behalen van zijn VWO diploma in 1986 klaargestoomd voor een begin aan een universitaire opleiding in de natuurwetenschappen.

In september 1986 begon de auteur aan een studie Technische Wiskunde verzorgd door de Faculteit der Technische Wiskunde en Informatica aan de Technische Universiteit Delft. Deze studie heeft hij in vijf jaar afgerond, zodat hij in 1991 afstudeerde als wiskundig ingenieur. Zijn afstudeeropdracht werd begeleid door Prof.dr. G.J. Olsder en Dr. J.W. van der Woude, die destijds waren verbonden aan de vakgroep der Toegepaste Analyse. Het afstudeeronderwerp betrof een methode ter berekening van het stationaire gedrag van stochastische seriële blokkerende produktielijnen.

Kort volgend op het afstuderen van de auteur is hij onder supervisie van Prof.dr. G.J. Olsder in september 1991 begonnen aan een promotie-onderzoek en daarvoor als assistent in opleiding verbonden aan de Technische Universiteit Delft bij de Faculteit der Technische Wiskunde en Informatica. Het onderzoek betrof een academisch project met promovendi van verschillende Westeuropese universiteiten in samenwerking met een bekende lampenfabrikant. Het doel van het project was de optimalisatie van een specifiek produktieproces van autolampen. Tijdens deze periode als onderzoeker werden ook een aantal cursussen, verzorgd door het landelijk netwerk Systeem- en regeltheorie, gevolgd. Het onderzoek resulteerde in enkele publicaties. Gedurende het onderzoek werd werk gepresenteerd op conferenties in Brighton (Engeland, 1991) en Parijs (Frankrijk, 1995). Daarnaast heeft hij bij diverse academische gelegenheden in Nederland voordrachten gehouden over zijn werk (Benelux Meeting, 1992, Delft Control, 1993, DISC colloquium, 1995 en andere colloquia).

In de jaren 1994 tot en met 1996 is de auteur actief lid geweest van de AIO-commissie van Delft Control, een interfacultaire organisatie die zich bezighoudt met samenwerking van regeltechnisch onderzoek binnen de universiteit.

Het is de bedoeling naar aanleiding van dit proefschrift om de periode van onderzoek af te ronden met het verkrijgen van de doctorstitel na een verdediging die zal plaatshebben op vrijdag 20 december 1996 om 10.30 's ochtends.

Acknowledgements

I would like to thank everyone who contributed to my physical and psychological well-being during the past five years of research.

I would like to thank my supervisor Geert Jan Olsder for giving me the opportunity to do research during those years. I thank him also for allowing an unforgettable trip to France where he created the perfect conditions for the start of the relationship with my girlfriend Catherine. I thank him and colleague Jacob van der Woude for their support and supervision in the research project. I would like to thank my project-mates Heinz Falkus, Peter van Overschee, Ghassan Murad, Richard Hakvoort and Jobert Ludlage for their cooperation. I would also like to thank Mr K. Dankwart for his cooperation in the car lamp factory.

I would like to thank my roommate Hans Braker, who showed me during the first two years how an enthusiastic self-assured manager executes his job. I thank him for the nice times that we spent together being mates in our room at the university as well as mates in rooms in cyberspace. I also thank Han Adriaens and Erwin Mouthaan for being nice company as roommates during the last three years and I wish them success with their future research activities.

Among the colleagues I would like to thank Peter Sonneveld especially for his help with the proof of one of the properties in this thesis. I thank all the other colleagues for the (more or less spontaneous) interesting conversations that we had in the local coffee room and for the nice times that I spent with most of them elsewhere.

I would like to thank Tatiana for her assistance with the basic facilities and material and Eef and Marion for their support with the facilities in the field of hard and software.

I would like to thank the members of the promotion committee for their effort of carefully reading this thesis and for their cooperation in the promotion ceremony.

I would like to thank my best friends Arthur, Paul, Arno, Hans and Vincent for being there in good and in bad times and I wish that we can continue friendship like this for the rest of our lives.

I am most grateful to my family: my mother, father and brother. They initiated my present physical and psychological being. They taught me the fundamentals of life each in their

own personal way and formed a big part of my present character with its good and bad qualities. Although they have produced a far from perfect human being, I think they have succeeded a great deal in creating the best circumstances for a happy life by supporting good education which is the essential key to a free mind. I am among the lucky ones with such a family who has always given me an unconditional support and who encouraged me to continue during the most difficult times in my life, especially the past five years of research.

Special feelings are for my girlfriend Catherine who I met during this period of research and whose strong personality helped me to gain self-confidence and feel new purpose in a future life.

List of Figures

List of Tables