Integrated Models for Performance Analysis and Optimization of Queueing-Inventory Systems in Logistic Networks

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät Mathematik, Informatik und Naturwissenschaften
Fachbereich Mathematik
der Universität Hamburg

vorgelegt
von
Sonja Otten

Hamburg
2017

Tag der Disputation: 31. Januar 2018

Folgende Gutachter empfehlen die Annahme der Dissertation:
Prof. Dr. Hans Daduna
Prof. Dr. Holger Drees

Vorsitzender der Prüfungskommission:
Prof. Dr. Jens Struckmeier

Danksagung

An dieser Stelle möchte ich mich bei allen herzlich bedanken, die mich bei der Erstellung dieser Arbeit unterstützt haben.

Mein großer Dank gilt meinen Eltern Gerda und Wolfgang Otten, die es mir ermöglicht haben zu studieren und zu promovieren. Ich bedanke mich ganz herzlich für ihre Unterstützung und ihren Glauben an mich nicht nur während der Promotion, sondern auch während des Studiums.

Ganz besonders danke ich Prof. Dr. Hans Daduna. Seit der Vorlesung zu Stochastischen Prozessen während meines Bachelorstudiums hat er die Begeisterung für die Warteschlangentheorie in mir geweckt. Er gab mir die Möglichkeit, mich mit einem interessanten Forschungsthema auseinanderzusetzen, und hat mich über den gesamten Zeitraum dieser Arbeit hervorragend betreut. Seine vielseitige und außerordentliche Unterstützung hat maßgeblich zum Gelingen dieser Arbeit beigetragen. Es war und ist mir eine große Freude, mit ihm gemeinsam zu forschen.

Ein herzlicher Dank geht an Prof. Dr. Holger Drees, dass er mir ermöglicht hat, als wissenschaftliche Mitarbeiterin im Bereich Mathematische Statistik und Stochastische Prozesse der Universität Hamburg zu arbeiten, um dort mein Promotionsprojekt umzusetzen. Des Weiteren danke ich ihm für seine Unterstützung und die Erstellung des Gutachtens.

Den anonymen Gutachtern danke ich für ihre hilfreichen Kommentare zu den Teilen dieser Arbeit, die wir bei Fachzeitschriften eingereicht haben.

Ein großes Dankeschön gilt dem gesamten Bereich Mathematische Statistik und Stochastische Prozesse. Allen jetzigen und ehemaligen Mitarbeiterinnen und Mitarbeitern danke ich für ihre Hilfsbereitschaft und die vielen anregenden Diskussionen. Sie alle haben in sowohl fachlichen als auch privaten Gesprächen dafür gesorgt, dass das Arbeitsumfeld stets angenehm und freundschaftlich war.

Insbesondere bedanke ich mich bei meinem Bürokollegen und Mitautoren Dr. Ruslan Krenzler für die hilfsbereite und freundschaftliche Arbeitsatmosphäre sowie die exzellente Kooperation. Seine wertvollen Hinweise und Ideen sowie seine konstruktive Kritik haben die Arbeit bereichert.

Des Weiteren danke ich ganz herzlich Dr. Victoria Magath, Dr. Torge Schmidt, Miran Knežević, Daniel Ebel und meiner Mutter für die Mühe, die sie sich beim Korrekturlesen dieser Arbeit gemacht haben.

Außerdem bedanke ich mich bei vielen Freundinnen und Freunden für ihre liebe Unterstützung, ihre Aufmunterungen und ihre entgegengebrachte Nachsicht.

Meinem Freund Dr. Karsten Kruse danke ich von ganzem Herzen für seine hilfreiche Unterstützung und sein Verständnis bei der Anfertigung dieser Arbeit. Er hat nicht nur unermüdlich Kapitel für Kapitel Korrektur gelesen, sondern auch mit fachlichen Diskussionen und wertvollen Ideen die Arbeit bereichert.

Contents

Introduction

Motivation

Today's production processes and production systems usually are large systems of interacting components, and the components are typically of very different nature, e.g. production centres, logistic and transport units, inventories, etc. A supply chain represents a "(...) network of organizations that are involved, through upstream and downstream linkages, in the different processes and activities that produce value in the form of products and services in the handy of the ultimate consumer" [Chr98, p. 15]. An example of a supply chain, which consists of customers, a production system, an inventory and a supplier, is presented in Figure 0.0.1. Understanding the functioning of these systems is an important issue and there is need for insight in the structure of these complex systems with strongly interacting subsystems. As can be seen from the recent literature, there is much research in the field of supply chains, but as it can be seen as well, structure theory for these complex systems is in a very premature status.

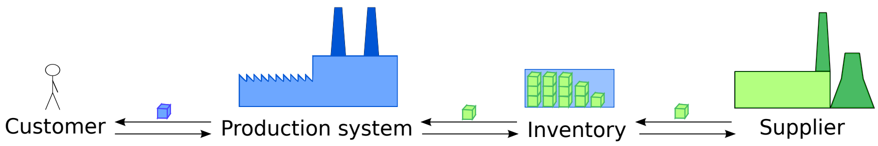


Figure 0.0.1.: Supply chain

Production processes are usually investigated using models and methods from queueing theory. Control of warehouses and their optimization rely on models and methods from inventory theory. Both theories are fields of Operations Research (OR), but they comprise quite different methodologies and techniques. In classical OR queueing and inventory theory are considered as disjoint research areas. On the other side, the emergence of complex supply chains ($\equiv$ production-inventory networks) calls for integrated production-inventory models as well as adapted techniques and evaluation tools. Such integrated approaches to model production-inventory systems have been developed over the last decade and it turned out that the problem of determining e.g. steady state distributions of the systems results in either large simulation experiments or in using heuristic decomposition-aggregation methods or in solving the global balance equations numerically.

In Operations Research and applied mathematics — especially applied probability — there exist well established theories for the components of the supply chains and the production systems. They are connected with, for example, queueing theory, inventory theory and transport theory. These theories provide structural characteristics, performance metrics, conditions for stabilization, and so on, which are useful in running such systems under optimal conditions. As an example: One of the most important tasks in business is inventory management, whereby the fundamental problem can be described by two questions (cf. [BCST09, p. 3]): “When should an order be placed?” and “How much should be ordered?”. To answer such questions we need the support of inventory theory, and indeed this theory provides answers at least for small inventories.

In this thesis both — queueing theory and inventory control — are methodologically relevant, in particular, integrated production-inventory models. Over the last decades research on queueing systems with attached inventory found much attention, often in connection with the research on integrated supply chain management. For a general review we refer to Krishnamoorthy et al. [KLM11]. Some additional articles can be found in [KS16a]. These articles are by no means complete. On page viii we describe the previously done research, which is relevant for our studies, in more detail.

The ***integrated models in the literature*** assume a continuous review structure and the supply chains of interest consist — in the fundamental version — of customers, a production system, an inventory and a supplier as shown in Figure 0.0.2.

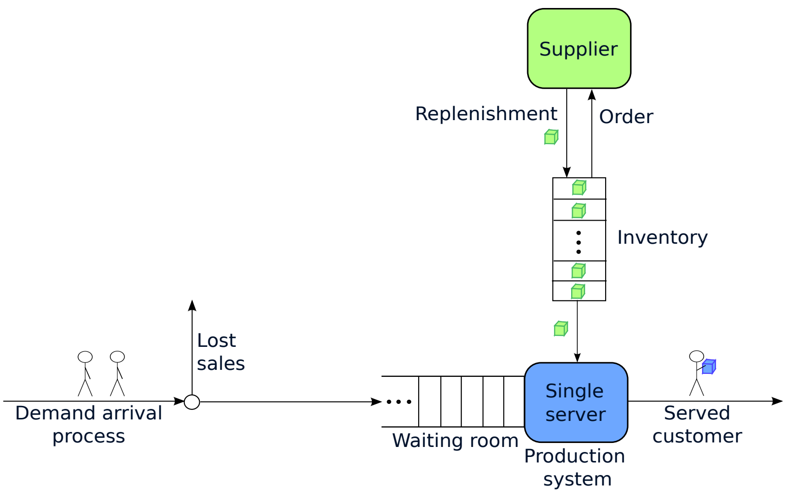


Figure 0.0.2.: Supply chain

The production system manufactures products according to customers’ demand on a make-to-order (MTO) basis[1], i.e. the manufacturing starts only after an order of a customer is received. According to a Poisson process indistinguishable customers arrive one by one at the production system and require service. There is a single server with waiting room under a first-come, first-served regime. Each customer needs exactly one

[1]The use of manufacturing terms in the literature is not consistent. We use the definition of Schneeweiß [Sch02, pp. 16f.]. He distinguishes only between make-to-order (MTO) and make-to-stock (MTS) as strategies for the production environment. In the literature, there are various variants of how many different strategies exist [Sin12, pp. 43f.]. For example, Stadtler et al. [SKM10, pp. 212-215] split the strategies further into assemble-to-order (ATO, also called capable-to-order).

item from the inventory for service. If the server is ready to serve a customer, who is at the head of the line and the inventory is not depleted, his service begins.

There are two extreme cases of customers' reaction in the situation that inventory is depleted when demand arrives (cf. [SPP98, p. 234]): Backordering, which means that customers are willing to wait for their demands to be fulfilled, and lost sales, i.e. demands that occur when inventory is empty are lost. In this thesis, we focus on lost sales models, like for example on a model depicted in Figure 0.0.2.

In these models, a served customer departs from the production system immediately and the associated item is removed from the inventory. It is assumed that the transportation time between the production system and the inventory is negligible. An outside supplier replenishes raw material in the inventory according to a continuous review replenishment policy. At each decision epoch, it is determined according to a prescribed replenishment policy whether a replenishment order is placed or not, and how many items are ordered.

In this thesis, we consider the following continuous review replenishment policies, whereby we focus on the base stock policy.

- *Base stock policy:*
  Each unit taken from the inventory results in a direct order for one unit sent to the supplier. This means, if a served customer departs from the system, an order of the consumed raw material is placed at the supplier at this time instant. The local base stock level $b \geq 1$ is the maximal size of the inventory. Note that there can be more than one outstanding order.
  An equivalent definition can, for example, be found in [HS00, p. 65].

- (r, Q)*-policy:*
  If the on-hand inventory falls down to a prefixed value $r \geq 0$, a replenishment order is placed instantaneously. The size of the order is fixed to $Q < \infty$ units of raw material. We assume that $r < Q$ (this "(...) ensures that there is no perpetual shortage" [LFW14, p. 1545]) and that there is at most one outstanding order. The maximal size of the inventory is $r + Q$.
  Equivalent definitions can, for example, be found in [SSD$^+$06, p. 63], [HS00, p. 65] and [SPP98, pp. 237f.]. Furthermore, Silver et al. [SPP98, pp. 237f.] give a brief discussion of the advantages and disadvantages for (r, Q)-policy.

- (r, S)*-policy:*
  If the size of the local inventory is less than or equal to the reorder level $r \geq 0$, a replenishment order is placed instantaneously. With each replenishment the local inventory level is restocked to exactly $S < \infty$. The maximal size of the inventory is S. We assume that $0 < S$ and that there is at most one outstanding order ($r < S$). Equivalent definitions can, for example, be found in [SSD$^+$06, p. 65], [BS01, p. 431] and [SPP98, pp. 238f.]. Furthermore, a brief discussion of the advantages and disadvantages can be found in [SPP98, pp. 238f.] for (r, S)-policy.

The models in the literature under investigation differ in service time distribution, lead time distribution, waiting room size, inventory capacity, replenishment policy and the costs which originate from the queueing of customers and from holding inventory. Furthermore, a distinction is made between the lost sales and the backordering case.

Literature review

For a general review we refer to Krishnamoorthy et al. [KLM11], which is by no means complete. Some additional articles can be found in [KS16a]. In the following, we describe the previous research, which is relevant for our studies, in more detail.

The first intensive study on a queueing-inventory model is conducted by Sigman and Simchi-Levi [SSL92]. They use an approximation procedure to find performance descriptions for an $M/G/1$ queue with limited inventory.

In a sequence of papers, Berman and his coauthors investigate the behaviour of production systems with an attached inventory. They define a Markovian system process and use classical optimization methods to find the optimal control strategy of the inventory.

In [BK99], Berman and Kim study queueing-inventory systems with Poisson arrivals, exponentially distributed service times and zero lead time under backordering with an infinite waiting room. The authors prove that the optimal replenishment policy does not place an order when the inventory level is positive; it places an order only when the inventory level drops to zero and the queue length exceeds some threshold value. They also model the case in which the waiting room is finite and customers, who arrive when the queue is full, are lost.

In another paper, Berman and Kim [BK01] extend their earlier model with the infinite waiting room and allows exponential or Erlang lead times for replenishment. For known order size Q, the optimal policy minimises the expected discounted costs and the average costs. They find out that the optimal ordering policy has a monotonic threshold structure.

The model in [BK04] can be viewed as an extension of the second paper [BK01] in the sense that it is assumed that a revenue is generated upon the service. They identify the optimal replenishment policy which maximizes the system profit.

Berman and Sapna [BS00] analyse queueing-inventory systems with Poisson arrivals, general service times and zero lead time under backordering. The size of the waiting room is finite and arriving customers are lost during the time the queue is full. They compute the steady state probabilities. Furthermore, the optimal value of the maximum allowable inventory size, which minimises the long-run-expected cost rate, is obtained and some performance measures are determined. Various examples of service time distributions (exponential, Erlang, constant) and optimal values for maximum inventory in each of these cases are also presented. Moreover, the authors consider the infinite waiting space case.

In another paper [BS01], Berman and Sapna investigate a system with Poisson arrivals, exponentially distributed service times and lead times under backordering. The size of the waiting room is finite and arriving customers are lost during the time the queue is full. They prove the existence of a stationary optimal policy. For given values of maximum inventory and reorder levels, they determine the service rates such that the long-run expected cost rate is minimised.

He and his coauthors [HJB02a] analyse $M/M/1/\infty$ production-inventory systems with zero lead time and backordering. They explore the structure of the optimal replenishment policy which minimises the average total cost per product.

In another paper [HJB02b], the authors study $M/PH/1/\infty$ production-inventory systems with Erlang distributed lead times and backordering. They quantify the value of information used in inventory control.

Schwarz and her coauthors [SSD$^+$06] investigate $M/M/1$ systems with inventory management, exponentially distributed lead times and lost sales. They consider order replenishment policies with a fixed reorder point and a general randomized order size as well as a deterministic order size. Further, they distinguish between an infinite and a finite waiting room. They derive stationary distributions of joint queue length and inventory processes in explicit product form and calculate performance measures of the respective systems.

Schwarz and Daduna [SD06] study $M/M/1/\infty$ systems with inventory management, exponentially distributed lead times and backordering. They concentrate on the case of $(0, Q)$-policy with and without an additional threshold. They calculate respectively approximate performance measures and derive optimality conditions under the different order policies.

Saffari and his coauthors [SHH11] provide an extension of Schwarz et al. [SSD$^+$06]. They prove that the $M/M/1/\infty$ system with inventory under (r, Q)-policy with hyperexponential lead times (i.e. mixtures of exponential distributions) has a product form distribution. The resulting distribution is employed to compute performance measures of the system.

Saffari and Haji [SH09] study a two-echelon supply chain which consists of a retailer and a supplier. Demands arrive according to a Poisson process at the retailer, who uses a base stock policy. The supplier follows an (r, Q)-policy and the service and replenishment lead times are exponentially distributed. When the supplier has no on-hand inventory, arriving demands from the retailer are lost. They calculate long-term performance measures of the system to find the optimal order size.

Haji and his coauthors [HHS11, HSH11] consider a two-echelon supply chain where the supplier is a service system with an attached inventory and both supplier and retailer apply a base stock policy. Demands arrive to the retailer according to a Poisson process. During the time that the supplier has no on-hand inventory, arriving demands are lost to the supplier and the retailer buys products from another source with zero lead time and with additional cost. Service times and replenishment lead times of the supplier's system are exponentially distributed. They derive the stationary distribution of joint queue length and on-hand inventory of the supplier and show that it is of product form. Furthermore, they derive the total expected system cost per unit time.

Saffari and his coauthors [SAH13] investigate $M/M/1/\infty$ systems with inventory under (r, Q)-policy and with lost sales. They derive the stationary distributions of the joint queue length and on-hand inventory when lead times are random and with various distributions. Furthermore, they formulate long-run average performance measures and cost functions in some numerical examples. [SHH11] and [SAH13] are (slightly) generalized (removing restrictions) in [Kre16] and [KD15].

Krishnamoorthy and Viswanath [KV11] are the first who report work on production inventory system with positive service time. The time for producing each item follows a Markovian production scheme. The customer arrival process follows a Markovian arrival process and the service time of each customer has a phase-type distribution. They investigate the stability of the system and compute several measures of system performance.

Krishnamoorthy and Viswanath study in [KV13] production-inventory systems with (r, S)-policy, positive service time and lost sales. They derive the joint stationary distribution in explicit product form. They develop a technique where the steady state vector of the classical $M/M/1$ system and the steady state vector of a production-inventory system, where the service is instantaneous and no backlogs are allowed, are combined. They apply their technique to the models discussed in [SSD+06].

Krishnamoorthy, Manikandan and Shajin analyse in [KMD15] an $M/M/c$ queueing-inventory system with positive service time and (r, Q)-policy. The required item is either provided after service with probability γ or else is not provided at the end of a service. For the case of two servers they obtain the steady state distribution in product form. For the case of more than two servers they do not have an analytical solution and analyse this case by an algorithmic approach. Furthermore, they derive an explicit expression for the stability condition and some conditional distributions. Moreover, they obtain several measures of system performance.

Krishnamoorthy, Shajin and Lakshmy study in [KSL16] a supply chain with one production centre and one distribution centre. Stocks are kept in both, the production centre to satisfy customers' demands and the distribution centre to satisfy demands from the distribution centre. The inventory at the distribution centre is controlled by an (r, Q)-policy. The production centre adopts an (sQ, KQ)-policy. The service time at the distribution centre is exponentially distributed and the lead time follows an exponential distribution. They derive the joint stationary distribution of the system in explicit product form.

Krishnamoorthy and Shajin analyse in [KS16b] an $M/M/1$ retrial queue with an inventory system and lost sales. The inventory is controlled by an (r, S)-policy and the replenishment lead time is exponentially distributed. Whenever the server is idle, arriving customers enter directly to an orbit. They derive the joint stationary distribution of the queue length and the on-hand inventory in explicit product form.

Thesis structure

We consider more complex supply chains than those described in the literature review above. In particular, we consider networks of production-inventory systems as shown in Figure 0.0.3. More precisely, in Part I production-inventory systems at several locations are connected by a supplier. Demand of customers arrives at each production system according to a Poisson process and is lost if the local inventory is depleted. To satisfy a customer's demand a server at the production system takes exactly one unit of raw material from the associated local inventory. The supplier manufactures raw material to replenish the local inventories, which are controlled by a continuous review base stock policy.

Chapter 2 to Chapter 4 are devoted to the research of the network's behaviour, where the supplier consists only of a single server and replenishes the inventories at all locations. The items of raw material are indistinguishable (exchangeable).

In Chapter 2, we investigate this model, that we consider to be the basic model.

In Chapter 3, we analyse an extension, where routing of items depends on the on-hand inventory at the locations (with the aim to obtain "load balancing"). The systems under

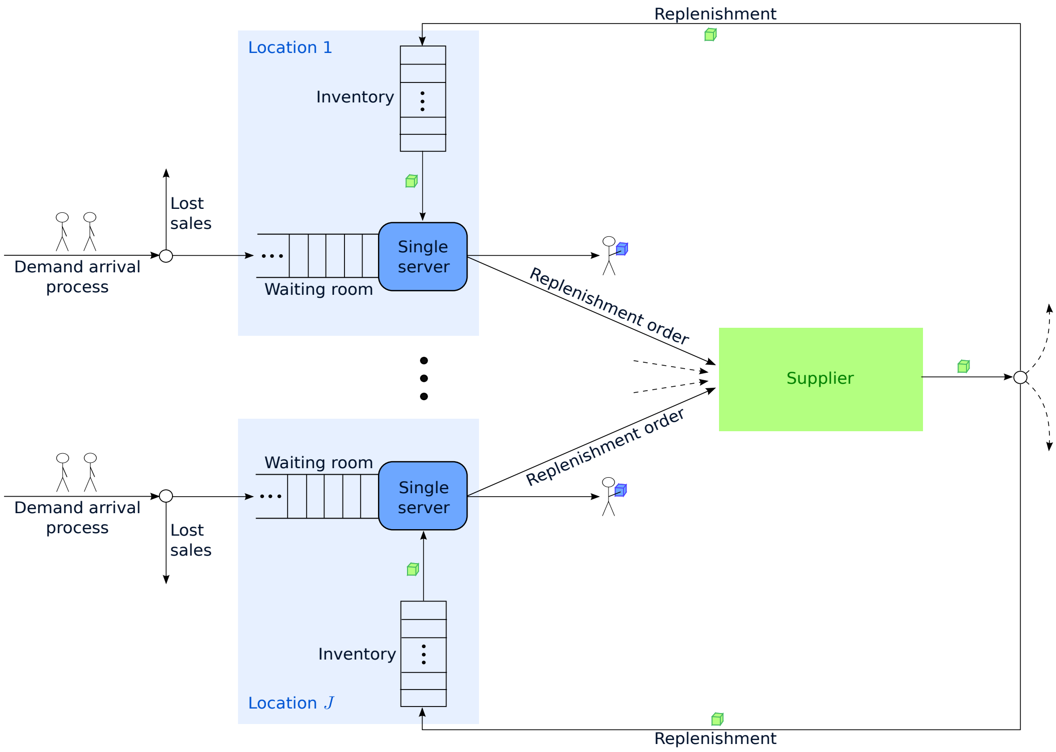


Figure 0.0.3.: Network of production-inventory systems

investigation differ with respect to the load balancing policy.

In Chapter 4, we study the basic model with perishable items, since in certain types of inventories the items either perish, deteriorate or become obsolete. This model is a special case of queueing systems in a random environment which we have introduced in Appendix D.1.

Chapter 5 to Chapter 7 are devoted to the research of the network's behaviour of more complex models, where the finished items are delivered exactly to the locations where the orders were generated, i.e. they are not exchangeable. These models can be classified as a "multi-product system".

In Chapter 5, we investigate this model, where the supplier is a complex network. This model is considered as basic model as well.

In Chapter 6, we look at the aggregation of the supplier network. We can substitute the complex supplier network by only one node — a supplier who consists of a symmetric server. The symmetric server enables to deal with non-exponential type-dependent service time distribution for different order types.

In Chapter 2 to Chapter 6, we focus on base stock policies. Nevertheless, in classic inventory theory several replenishment policies are considered. Hence, we investigate the (r, S)-policy in Chapter 7. The systems under investigation differ with respect to the reorder level and the number of locations and workstations.

In this thesis, we study the stability behaviour of these integrated production-inventory systems. For the most of these integrated production-inventory systems the obtained steady state is of so-called "product form", which reveals a certain decoupling of the components of the system for long time behaviour. The simple structure of this steady state allows to apply "product form calculus", a widely used tool, which provides access to easy performance evaluation procedures. More details about the art of product form modelling can be found in Section 1.1. Moreover, computational algorithms to calculate important performance measures are developed and (with the help of these) cost analysis for these systems is demonstrated.

Up to now, one of the key assumptions of production-inventory models in literature is that customers are indistinguishable. In practice, however, customers have different characteristics and/or priorities, which leads to systems where this assumption does not hold. Therefore, Part II is devoted to the study of multiple customer classes with different priorities. The research is dedicated to production-inventory systems with two classes of customers and inventory management under lost sales where the customers' arrivals are regulated by a flexible admission control as shown in Figure 0.0.4. We have investigated the (r, Q)-policy in Chapter 10 and the base stock policy in Chapter 11. We derive some structural properties of the steady state distribution which provide insights into the equilibrium behaviour of the systems. Moreover, the existence of a stationary distribution is investigated. Furthermore, we consider for these systems the case of zero service time, which is the version of our model in the classical inventory theory.

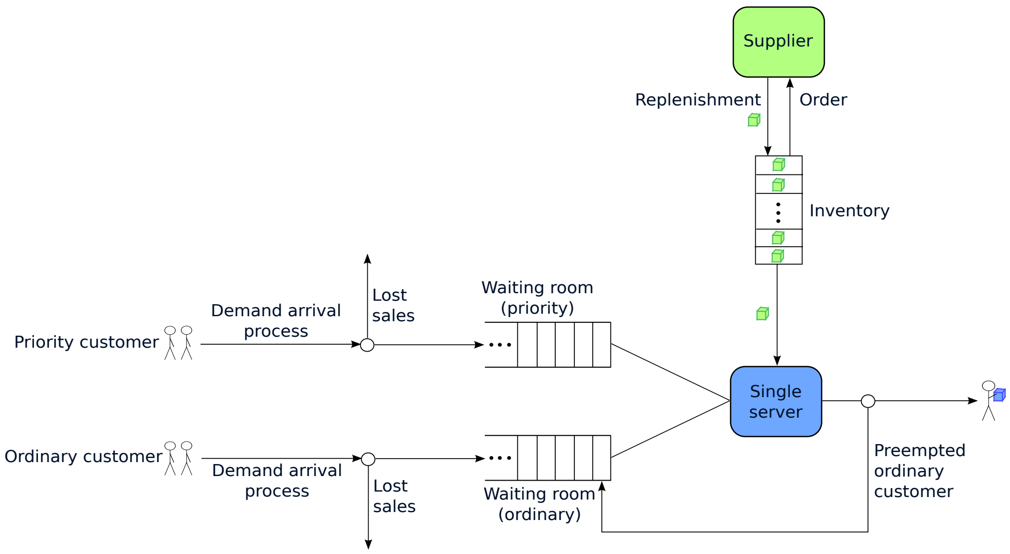


Figure 0.0.4.: The production-inventory system with distinguishable customers

Notation and preliminaries

$\mathbb{N} := \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$, $\mathbb{R}_0^+ := [0, \infty)$, $\mathbb{R}^+ := (0, \infty)$. $\mathbb{B}$ are the Borel sets of $\mathbb{R}$.

A value is said to be *positive* if it is greater than zero and a value is said to be *negative* if it is less than zero. We call a value *non-positive* if it is less than or equal to zero. We call a value *non-negative* if it is greater than or equal to zero.

The vector $\mathbf{0}$ is a row vector of appropriate size with all entries equal to 0. The vector $\mathbf{e}$ is a column vector of appropriate size with all entries equal to 1. The vector $\mathbf{e}_i = (0, \ldots, 0, \underbrace{1}_{i-\text{th element}}, 0, \ldots, 0)$ is a vector of appropriate dimension.

The notation $\subset$ between sets means "subset or equal" and $\subsetneq$ means "proper subset". We write $C = A \uplus B$ to emphasize that C is the union of disjoint sets A and B. For a set A we denote by $|A|$ the number of elements in A.

The notation $x \approx y$ means x is approximately equal to y.

$1_{\{expression\}}$ is the indicator function which is 1 if *expression* is true and 0 otherwise.

Empty sums are 0, and empty products are 1.

For $k > 1$ and $m, \ell \in \{1, \ldots, k\}$ we call for $m \leq \ell$ the sequence $m, m+1, \ldots, \ell$ a list. If $m > \ell$, the list $m, \ldots, \ell$ is the empty list.

For $x > 0$ we define $\frac{1}{0} := \infty$, $\frac{0}{0} := 0$ and $\infty \cdot 0 := 0 \cdot \infty := 0$.

We call a *generator* a matrix $M \in \mathbb{R}^{K \times K}$ with countable index set K, whose all off-diagonal elements are non-negative and all row sums are equal to zero. By definition this implies that the diagonal elements are finite.

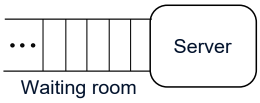
We call a matrix $M \in [0, 1]^{K \times K}$ with countable index set K *stochastic* if the row sums are one.

We call a matrix $M \in [0, 1]^{K \times K}$ with countable index set K *substochastic* if the row sums are less than or equal to one.

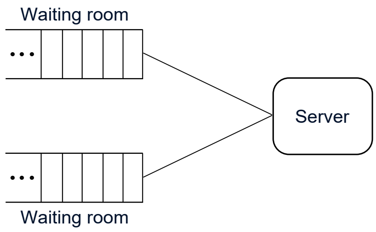
Throughout this thesis it is assumed that all random variables are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Furthermore, by Markov process we mean time-homogeneous continuous-time strong Markov process with discrete state space ($\equiv$ Markov jump process). Without further mentioning all Markov processes are assumed to be regular and have cadlag paths, i.e. each path of a process is right-continuous and has left limits everywhere. We call a Markov process regular if it is non-explosive (i.e. the sequence of jump times of the process diverges almost surely), its transition intensity matrix is conservative (i.e. row sums are 0) and stable (i.e. all diagonal elements of the transition intensity matrix are finite).

In this thesis, different chapters consider different models, which can be read independently of each other. Because of this, some remarks become repetitive.

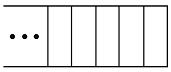
In diagrams, see for example Figure 0.0.5, we use rounded rectangles to represent servers of the queues.



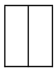
(a) Server with waiting room



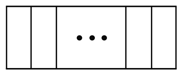
(b) Server with two waiting rooms



(c) Waiting room with infinite number of waiting places



(d) Waiting room with two waiting places



(e) Waiting room with finite but unknown or large number of waiting places

Figure 0.0.5.: Symbolic representation of the queues

The use of manufacturing terms in industry and in the literature is far from standardized as mentioned by Hopp and Spearman [HS00, p. 215]. Hence, we will define our terms in the following and caution the reader that the same terms can be used differently in other sources.

- A **queueing system** consists of a server (single server or multiple server) and waiting room(s). A server without waiting room is also called a queueing system.
- A **supply chain** represents a "(...) network of organizations that are involved, through upstream and downstream linkages, in the different processes and activities that produce value in the form of products and services in the handy of the ultimate consumer" [Chr98, p. 15]. The terms **production-inventory system**, **production-inventory-replenishment system** and **queueing-inventory system** as well as **integrated queueing-inventory system** are synonymous with supply chain in this thesis.
- A **location** consists of a production system with attached inventory.
- A **production system** is modeled as a queueing system.
- An **inventory** is replenished by the supplier (network) with raw material.

- The **on-hand inventory** is the size of the inventory, i.e. the number of items of raw material which are on stock or in production.
- **Raw materials** are "(...) items purchased from suppliers to be used as inputs into the production process. They will be modified or transformed into finished goods" [Ter88, p. 4]. Hopp and Spearman mentioned that raw materials are "(...) components, subassemblies, or materials that are purchased from outside the plant and used in the fabrication/assembly processes inside the plant" [HS00, p. 582].
- **Item** is the abbreviation for "item of raw material". In the literature synonymous with items are parts, components, subassemblies, assemblies.
- A **supplier (network)** consists of workstations, it manufactures raw material to be forwarded to the inventory.
- A **workstation** is modeled as a queueing system at a supplier.
- **Orders** are the units at the workstations of the supplier (network). In the literature it is often called work in process (WIP) (cf. [HS00, p. 582]).

Part I.

Networks of production-inventory systems

1. Introduction

Integrated approaches to model production-inventory systems have been developed over the last decade and it turned out that the problem of determining e.g. steady states of the systems usually results in either large simulation experiments or in using heuristic decomposition-aggregation methods or in solving the global balance equations numerically.

We consider a network of production-inventory systems as shown in Figure 1.0.1, which consists of parallel production systems (single servers) at different locations each with an attached local inventory, and a supplier, which produces raw material (discrete units) to replenish the local inventories.

In Chapter 2 to Chapter 4 the supplier consists only of a single server and replenishes the inventories at all production locations. The items of raw material are indistinguishable (exchangeable). Chapter 5 to Chapter 7 is devoted to the research of the network's behaviour of a more complex model, where the finished items are delivered exactly to the locations where the orders were generated, i.e. they are not exchangeable. It can be classified as a "multi-product system".

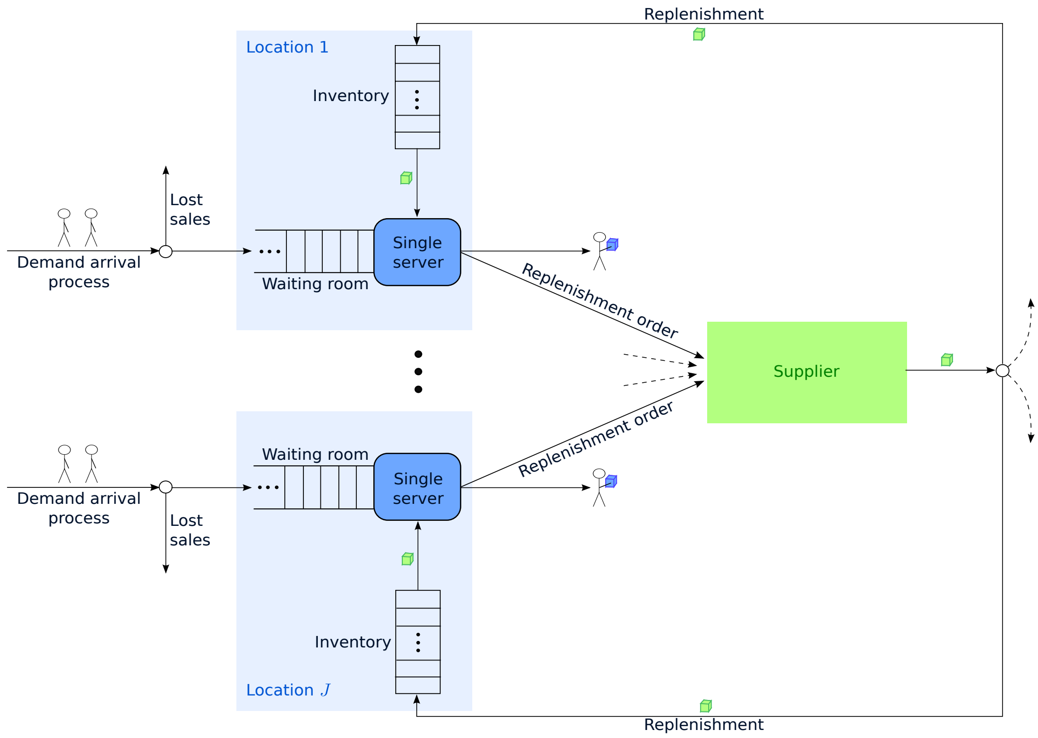


Figure 1.0.1.: Production-inventory system

Each arriving customer at the locations (production systems) initiates a production process that requires one item of raw material from the attached local inventory. Production at a location can start only when raw material at the local inventory is available. Newly arriving customers who see the inventory depleted will not enter the location ("lost sales"). Otherwise, new arrivals at a location enter the queue there and will wait until the previous customers' processing is finished. If no raw material is available for customers in the queue to start production, these customers will wait until raw material arrives at the local inventory. Consequently, there can be more customers waiting than the on-hand inventory level. All local inventories are replenished by the supplier network according to a continuous review base stock policy: Taking an item from the associated local inventory results in an order sent to the supplier network. Production of raw material only starts when there is a replenishment order.

Although we describe our systems in terms of production and manufacturing, there are other applications where our model applies, e.g. distributed retail systems where customers' demand has to be satisfied from the local inventories and delivering the goods to the customers needs a non-negligible amount of time; the replenishment for the local retail stations is provided by a production network. Another setting is a distributed set of repair stations where spare parts are needed to repair the brought-in items which are held in local inventories. Production of the needed spare parts and sending them to the repair stations is again due to a production network.

Several integrated production-inventory models are the focus of our present research. Our methodological approach constitutes an alternative to simulations and/or heuristic decomposition-aggregation techniques. We develop Markovian stochastic models of the production-inventory systems, which is smooth enough to be amenable to solving the steady state problem explicitly with closed form expressions for the stationary distribution. Moreover, for most of the models it turns out that the obtained steady state distribution is of a form which is well-known in pure queueing theory: We come up with a product form equilibrium for the integrated queueing-inventory system. This product form structure of the joint stationary distribution is often characterised as the global process being "separable", and is interpreted as "the components of the system decouple asymptotically and in equilibrium". Clearly, separability is an important (but rather rare) property of complex systems.
The simple structure of this steady state allows to apply "product form calculus", a widely used tool, which provides access to easy performance evaluation procedures. Moreover, computational algorithms to calculate important performance measures are developed and (with the help of these) cost analysis for these systems is demonstrated.

Different from the standard product form equilibria in queueing networks the steady state obtained for some integrated models is stratified. In the upper stratum, we obtain three vectors for production, inventory, and supplier network. In the lower stratum each of these vectors is composed of homogeneous coordinates. The product form inside the lower stratum resembles on one side (for the production subsystem) the independence structure of Jackson networks, and on the other side (for the inventory-replenishment subsystem) the conditional independence of Gordon-Newell networks. The inventory-replenishment subsystem is henceforth referred to as inventory subsystem as usual in inventory theory. If necessary, we explicitly point out to the difference.

1.1. The art of product form modelling: Separable networks

Parts of this section are taken from [OKD17].

The aim of product form modelling as a branch of queueing network theory is to construct easy to understand models for large systems with complicated structure. "Easy to understand" means that the main first-order performance characteristics of the network can be computed from the steady state distributions which are explicitly accessible. "Product form" refers to the observation that the steady state distribution of such models with a vector valued state process (e.g. the joint queue length process of a queueing network) is the product of the marginal steady state distributions (the queue lengths at the individual nodes of the network). For stable networks this means that in the long run and in stationary state the local behaviour of the nodes seem to decouple into independent or conditionally independent processes. Breakthroughs in the field of queueing network theory and its applications in various fields of operations research were the findings of Jackson [Jac57, Jac63] and Gordon and Newell [GN67], who discovered product form solutions of the global balance equations for classes of queueing networks. Their models are networks of exponential service stations and look rather simple with respect to the assumptions on the stochastic data underlying the networks' behaviour. In Appendix A.2, we summarize definitions and theorems on classical exponential networks. Nevertheless, it turned out that many real world systems exhibit astonishing robustness with respect to deviations from the structural and distributional assumptions that underlie the Jackson and Gordon-Newell networks. Subsequently product form models became popular in many fields of applications. A short review of experiences with modelling and performance analysis using product form techniques is Vernon's survey paper [Ver04] with additional references.

Nevertheless, product form modelling has to impose severe restrictions on the structure of the systems under consideration. Henderson [Hen90] discussed in detail: "When do we give up on product form solutions." But in that paper he presented a nice example of product form models for transmission protocols in telecommunications.

While the Jackson and Gordon-Newell networks were invented to model production networks (flow shops), an important subsequent application was modelling the ARPANET using Jackson networks by Kleinrock [Kle64]. This popularised product form models in computer science and enforced research on computer systems and computer and telecommunications networks. This lead Baskett, Chandy, Muntz, and Palacios [BCMP75] and Kelly [Kel76] to develop more complex product form models.

There are books available that deal with modelling, performance analysis and general network theory in the spirit of product form calculus, e.g. Kelly [Kel79], Walrand [Wal88], Serfozo [Ser99], Chao, Miyazawa, and Pinedo [CMP99] for networks in continuous time. Product form networks in discrete time are investigated in [CMP99] and in the books of Woodward [Woo94] and Daduna [Dad01a]. In addition, [CMP99] presents results from network theory where explicit steady state distributions are derived analogously, although the final results are no longer of product form in a strict sense, see for more information Henderson's discussion in [Hen90].

On the other side, there exist limitations when modifying the original Jackson and Gordon-Newell formalisms. Notoriously hard are two classes of models: (i) networks where nodes (servers) have finite waiting rooms which results in blocking phenomena, see Perros [Per90], Balsamo, De Nitto Persone, and Onvural [BDO01], and (ii) networks where the nodes (servers) are unreliable, break down and have to be repaired before servicing can continue, see Chakka and Mitrani [CM96].

Finite waiting rooms can be considered as intrinsic restrictions, breakdown due to environmental influences are external restrictions for the development of the queueing networks. Both of these restrictions occurred in some models developed during the last fifteen years by many researchers: A class of two-component hybrid systems which have a queueing component and a second component which is an attached inventory. From the viewpoint of the queue the restrictions imposed on the service process by the inventory are external, while from the integrated system these restrictions are intrinsic. We will be faced with both of these restrictions in our quest for product form steady states.

An important question is "Can we use our product form results to obtain simple product form bounds for the system with unknown non-product form stationary distribution?". This question is motivated by van Dijk and his coauthors (e.g. [Dij11b, Section 1,7, pp. 62f.], [Dij98, pp. 311ff.], [DK92], [DW89]). They show that a product form modification turns out to be quite fruitful to provide product form bounds for the throughput of a unsolvable ($\equiv$ unknown stationary distribution) queueing-inventory system. For example, van Dijk shows in [Dij11b, Section 1,7, pp. 62f.] a product form approximation for the simple but unsolvable tandem queue with finite waiting room at both stations. We will deal with separable approximation of non-separable systems in Section 4.3.3 in the model with perishable items in the inventory.

1.2. Related literature

Parts of this section are published in [OKD16].

Relevant for our research are queueing theory and inventory control, in particular integrated queueing-inventory models.

Literature on queueing theory is overwhelming, so we point only to the most relevant sources for our present investigation. Our production systems are classical $M/M/1/\infty$ queueing systems which constitute a network of parallel queues connected to the central supplier queue, cf. Kelly [Kel79] and Chao, Miyazawa, and Pinedo [CMP99] for general networks of queues.

Special queueing networks, which model multi-station maintenance and repair systems, are investigated by Ravid, Boxma, and Perry [RBP13] and Daduna [Dad90] and references therein. In these systems, circulating items are "exchangeable". This feature will occur in our model in Part I as well.

A study of queueing networks, which proceeds as we do in Section 2.4, is reported by van der Gaast et al. [GKAR12]. They describe in a first step a complicated network with finite buffers and deterministic routing and replace this in a second step by an analytically tractable network with random routing. Similar to our results they obtain closed-form expressions for the steady state distribution of the substitute network.

Literature on inventory theory is, similar to that on queueing theory, overwhelming, so we only point to some references closely related to our investigations. We mention that there are two extreme cases of arriving customers' reactions in the situation that inventory is depleted when demand arrives (cf. Silver, Pyke, and Peterson [SPP98]): Either backordering, which means that customers are willing to wait for their demands to be fulfilled, or lost sales, which means that demand is lost when no inventory is available on hand.

In classical inventory theory it is common to assume that excess demand is backordered (Silver, Pyke, and Peterson [SPP98], Zipkin [Zip00, p. 40], Axsäter [Axs00]). However, studies by Gruen, Corsten, and Bharadwaj [GCB02] and Verhoef and Sloot [VS06] analyse customers' behaviour in practice and show that in many retail settings most of the original demand can be considered to be lost in case of a stockout.

For an overview of the literature on systems with lost sales we refer to Bijvank and Vis [BV11]. They present a classification scheme for the replenishment policies most often applied in literature and practice, and they review the proposed replenishment policies, including the base stock policy. According to van Donselaar and Broekmeulen [DB13] "Their literature review confirms that there are only a limited number of papers dealing with lost sales systems and the vast majority of these papers make simplifying assumptions to make them analytically tractable."

Rubio and Wein [RW96] and Zazanis [Zaz94] investigated classical single item and multi-item inventory systems. Similar to our approach they used methods and models from queueing theory to evaluate the performance of base stock control policies in complex situations.

Reed and Zhang [RZ17] study a single item inventory system under a base stock policy with backordering and a supplier who consists of a multi-server production system. Their aim is to minimise a combination of capacity, inventory and backordering costs. They develop a square-root rule for the joint decision. Furthermore, they justify the rule analytically in a many-server queue asymptotic framework.

Because we consider queueing-inventory systems where inventories are controlled by base stock policies, we mention here that Tempelmeier [Tem05, p. 84] argued that base stock control is economically reasonable if the order quantity is limited because of technical reasons.

The base stock policy is "(...) more suitable for item with low demand, including the case of most spare parts" [RM11, p. 661].

Morse [Mor58, p. 139] investigated (pure) inventory systems that operate under a base stock policy. He gives a very simple example where the concept "re-order for each item sold" is useful: Items in inventory are bulky, and expensive (automobiles or TV sets[1]). He uses queueing theory to model the inventory systems, analogously to [RZ17], etc.

[1] The paper is from 1958.

Literature on integrated queueing-inventory models (i.e. queueing theory in combination with inventory theory) with non-zero lead times is surveyed by Krishnamoorthy, Lakshmy, and Manikandan [KLM11]. They give a review on inventory models, where items are delivered to customers on a first-come, first-served basis and it requires a non negligible amount of time. This time to deliver an item can be considered as a service time associated with the arriving demand. Reducing our models to the simplest situation with only one production (or service) unit and one inventory leads to a model investigated there. Furthermore, for production-inventory systems with positive service time we refer to Krishnamoorthy and Viswanath [KV11, KV13].

Literature on the system extensions can be found in the respective subchapters. However, it is understood that the main literature from this section is relevant for the models in Chapter 2 to Chapter 7 as well.

Exchangeable items

2. Basic production-inventory model with base stock policy

Parts of this chapter are published in [OKD16].

2.1. Own contributions

We develop a Markov process model of a complex supply chain, which encompasses production systems at several locations with associated local inventories, and a central supplier. We derive stationary distributions of joint queue length and inventory processes in explicit product form. After performing a cost analysis, we find out that the global search for the vector of optimal base stock levels can be reduced to a set of independent optimization problems. The explicit form of the stationary distribution enables us to get additional structural insights, e.g. about monotonicity properties and stability conditions. We show that our model — with the send out procedure of the central supplier by a random selection scheme — can be seen as an approximation for a model, where the finished items are delivered exactly to the locations where the orders were generated (for more details see Chapter 5).

If we consider the production facilities (queues) at the locations as devices (servers) which deliver items from the inventory to incoming demand, needing non-negligible delivering time (as in the single-echelon inventory systems case described by Krishnamoorthy, Lakshmy, and Manikandan [KLM11]), our results extend their setting to a multi-dimensional system.

On the other hand, our work is an extension of the investigations of Rubio and Wein [RW96], Zazanis [Zaz94] and Reed and Zhang [RZ17] on inventory systems under base stock policy: In their models there is no production-to-order such that the time to satisfy customer demand is zero. Therefore, their model is a special case of our model when the service time is set to zero. We need to mention that the replenishment system can be more complex in the mentioned papers than our replenishment server in this section (more complex replenishment servers will be considered in Chapter 5).

2.2. Description of the model

The supply chain of interest is depicted in Figure 2.2.1.[1] We have a set of locations $\overline{J} := \{1, 2, \ldots, J\}$. Each of the locations consists of a production system with an attached inventory. The inventories are replenished by a single central supplier, which is referred to as workstation $J + 1$ and manufactures raw material for all locations. The items of raw material are indistinguishable (exchangeable).

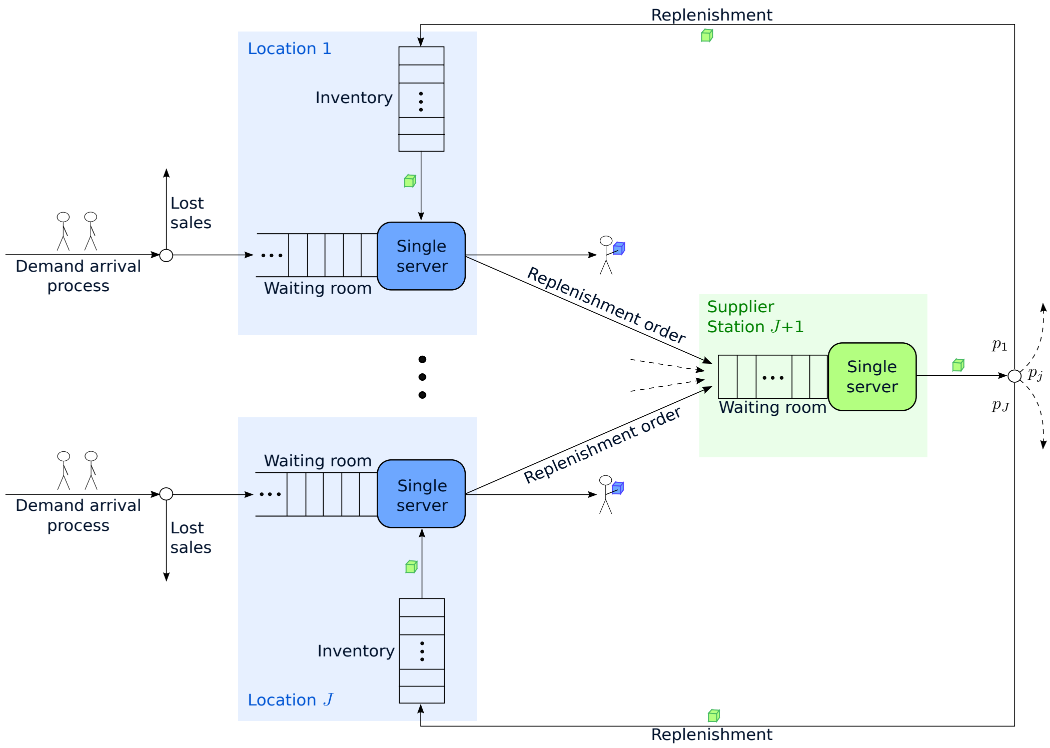


Figure 2.2.1.: Supply chain with base stock policy

Facilities in the supply chain. Each production system $j \in \overline{J}$ consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under a first-come, first-served (FCFS) regime. Customers arrive one by one at production system j according to a Poisson process with rate $\lambda_j > 0$ and require service. To satisfy a customer's demand the production system requires exactly one item of raw material, which is taken from the associated local inventory. When a new customer arrives at a location while the previous customer's order is not finished, this customer will wait. If the inventory is depleted at location j, the customers who are already waiting in line will wait, but new arriving customers at this location will decide not to join the queue and are lost ("local lost sales").

The service requests at the locations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of

[1] Figure 2.2.1 is the coloured version of the monochrome Figure 1 in our paper [OKD16, Figure 1].

the arrival streams. The service at location $j \in \overline{J}$ is provided with local queue-length-dependent intensity. If there are $n_j > 0$ customers present at location j either waiting or in service (if any) and if the inventory is not depleted, the service intensity is $\mu_j(n_j) > 0$. If the server is ready to serve a customer who is at the head of the line, and the inventory is not depleted, the service immediately starts. Otherwise, the service starts at the instant of time when the next replenishment arrives at the local inventory.

The inventory at location j is controlled by prescribing a local base stock level $b_j \geq 1$, which is the maximal size of the inventory at location j, we denote $\mathbf{b} := \big(b_j : j \in \overline{J}\big)$.

The central supplier (which is referred to as workstation $J+1$) consists of a single server (machine) and a waiting room under FCFS regime. At most $\sum_{j\in\overline{J}} b_j - 1$ replenishment orders are waiting at the central supplier. Service times at the central supplier are exponentially distributed with parameter $\nu > 0$.

All inter-arrival times at the locations and service times at the central supplier constitute an independent family of random variables.

Routing in the supply chain. A served customer departs from the system immediately after the service and the associated consumed raw material is removed from the inventory and an order for one item of the consumed raw material is placed at the central supplier ("base stock policy").

A finished item of raw material departs immediately from the central supplier and is sent to location $j \in \overline{J}$ with probability $p_j > 0$, independent of the network's history. $(p_j : j \in \overline{J})$ represents a predetermined delivering schedule with $\sum_{j\in\overline{J}} p_j = 1$. If the inventory is not full at location j (this means that the on-hand inventory level at location j is lower than the base stock level b_j), the item is added to the inventory at that location. Otherwise the item is added to the head of the queue of the central supplier, who will spend extra time on the already finished item and resend it to a new location $i \in \overline{J}$ according to the predetermined probabilities p_i, independent of the network's history.

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the central supplier and the local inventories are negligible.

Remark 2.2.1. The independence of the inter-arrival times and service times and the conditional independence of the routing in the supply chain is henceforth summarised as "usual independence assumptions".

Similar appropriate independence assumptions for the other models will be summarized by "usual independence assumptions" as well. Model-specific peculiarities will be mentioned separately, if necessary.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the size of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_{J+1}(t)$ we denote the number of replenishment orders at the central supplier at time $t \geq 0$ either waiting or in service (queue length)[2].

[2]The number of replenishment orders at the central supplier is denoted by $Y_{J+1}(t)$ in [OKD16].

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), \ldots, X_J(t), Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) \,:\, t \geq 0)\,.$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \left\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J,\, \mathbf{k} \in K\right\}$$

with

$$K := \{(k_1, \ldots, k_J, k_{J+1}) | 0 \leq k_j \leq b_j,\ j = 1, \ldots, J,\ k_{J+1} = \sum_{j=1}^{J}(b_j - k_j)\} \subset \mathbb{N}_0^{J+1}.$$

Note the redundancy in the state space: $W_{J+1}(t) = \sum_{j\in\overline{J}} b_j - \sum_{j\in\overline{J}} Y_j(t)$. We prefer to carry all information explicitly with because the dynamics of the system are easier visible.

Our aim is to analyse the long-run system behaviour and to minimise the long-run average costs.

Discussion of the modelling assumptions

We have imposed several simplifying assumptions on the production-inventory system to obtain explicit and simple-to-calculate performance metrics of the system, which give insights into its long-time and stationary behaviour. This enables a parametric and sensitivity analysis that is easy to perform.

First, the assumption of exponentially distributed inter-arrival and service times are standard in the literature and are the best first-order approximations. The locally state-dependent service rates are also common and give quite a bit of flexibility. The lead time is composed of the waiting time plus the production time at the central supplier. Therefore, it is more complex than exponential, constant or even zero lead times (which are often assumed in standard inventory literature). Zero lead times in our systems would result in almost trivial extensions of the queueing systems.

Second, we assume that the local base stock levels are positive (i.e. $b_j \geq 1$ at location j). This assumption can be made without loss of generality. Otherwise, all customers at location j would be lost, which is the same as excluding location j from the production-inventory system.

Third, the assumption of zero transportation times can be removed by inserting special (virtual) $M/G/\infty$ workstations into the network.

The fourth and most critical assumption from our point of view is the allocation of raw material from the central supplier to the production locations. We introduce a randomized decision scheme to select the target location, based on the "routing probabilities" p_j, $j = 1, \ldots, J$, and an additional acceptance-rejection rule. If the selected location j has a replenishment order outstanding, the item of raw material is sent to location j. Otherwise, the item of raw material remains in the machine of the central supplier for extra service after which the raw material is sent to a new location according to the same probabilities p_j (this is the same as discarding the item of raw material and placing a new replenishment order at the central supplier).

The latter assumption resembles some routing schemes from the literature, implemented in networks that are quite different from our model. It is well known that networks which encompass features like queues with buffers of finite size and/or with breakdowns of nodes have no simple explicit solution of the balance equations for the stationary distribution (see [Dij11a, Section 9.1, Section 9.4, Section 9.5]). There are two common strategies of rerouting to handle buffers of finite sizes (which could be applied in case of full inventories in our setting) in the literature:

1. "Skipping" principle: If a customer selects a node j, where the buffer is full, he only performs an imaginary jump to that node, spending no time there, but jumps onto a next node immediately according to the routing matrix and so on until he finds a free buffer place. This rerouting scheme is also known as "jump over protocol".

2. Blocking principle "repetitive service — random destination" (RS-RD): If a customer at node i selects a node j where the buffer is full, the customer stays at node i to obtain another service, after which the customer again selects a destination node according to the routing table and so on.

The skipping principle was introduced by Schassberger [Sch84] and later on was used e.g. in [Dij88], [Dij93] and in [DS96]. The RS-RD principle occurred as ALOHA-protocol e.g. in [Kle76, Section 5.11]. [SD04] discussed both principles and gave a short survey about the most prominent routing strategies in case of blocking.

In Section 2.4 we evaluate with the use of simulation whether the model with the abovementioned assumptions is a useful approximation for a more complex system where replenishment orders at the central supplier are dealt with in a FCFS order.

Remark 2.2.2. Appropriate discussions can be done for the extended models in Chapter 3 and Chapter 4. However, we will only discuss new assumptions in the extended models.

2.3. Limiting and stationary behaviour

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n},\mathbf{k}) \in E$:

$$\begin{aligned} q\left((\mathbf{n},\mathbf{k});(\mathbf{n}+\mathbf{e}_i,\mathbf{k})\right) &= \lambda_i \cdot 1_{\{k_i>0\}}, && i \in \overline{J}, \\ q\left((\mathbf{n},\mathbf{k});(\mathbf{n}-\mathbf{e}_i,\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\right) &= \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{k_i>0\}}, && i \in \overline{J}, \\ q\left((\mathbf{n},\mathbf{k});(\mathbf{n},\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\right) &= \nu p_i \cdot 1_{\{k_i<b_i\}}, && i \in \overline{J}. \end{aligned}$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

Proposition 2.3.1. *The strictly positive measure* $\mathbf{x} := (x\left(\mathbf{n},\mathbf{k}\right) : (\mathbf{n},\mathbf{k}) \in E)$ *with*

$$x\left(\mathbf{n},\mathbf{k}\right) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}\left(\mathbf{k}\right), \tag{2.3.1}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{2.3.2}$$

$$\widetilde{\theta}(\mathbf{k}) = \widetilde{\theta}(k_1, \ldots, k_J, k_{J+1}) = \prod_{j \in \overline{J}} \left(\frac{\nu p_j}{\lambda_j}\right)^{k_j}, \quad \mathbf{k} \in K, \tag{2.3.3}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z.

Remark 2.3.2. It has to be noted that k_{J+1} occurs only implicitly on the right side of (2.3.3). This hides a strong negative correlation of the coordinate processes $W_{J+1}(t)$ and $(Y_1(t), \ldots, Y_J(t))$ which is due to the state space restrictions.

Proof of Proposition 2.3.1. Note that $k_{J+1} > 0$ holds if $k_i < b_i$ for some $i \in \overline{J}$. Therefore, the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are:

$$\begin{aligned}
&x(\mathbf{n}, \mathbf{k}) \cdot \Big(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu p_i \cdot 1_{\{k_i < b_i\}}\Big) \\
= &\sum_{i \in \overline{J}} x(\mathbf{n} - \mathbf{e}_i, \mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} \\
&+ \sum_{i \in \overline{J}} x(\mathbf{n} + \mathbf{e}_i, \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \mu_i(n_i + 1) \cdot 1_{\{k_i < b_i\}} \\
&+ \sum_{i \in \overline{J}} x(\mathbf{n}, \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i \cdot 1_{\{k_i > 0\}}.
\end{aligned}$$

It has to be shown that the stationary measure (2.3.1) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

Substitution of (2.3.1) and (2.3.2) into the global balance equations directly leads to

$$\begin{aligned}
&\left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j)\right) \cdot \widetilde{\theta}(\mathbf{k}) \\
&\cdot \Big(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \cdot p_i \cdot 1_{\{k_i < b_i\}}\Big) \\
= &\sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J} \setminus \{i\}} \widetilde{\xi}_j(n_j)\right) \widetilde{\xi}_i(n_i - 1) \cdot \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} \\
&+ \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J} \setminus \{i\}} \widetilde{\xi}_j(n_j)\right) \cdot \widetilde{\xi}_i(n_i + 1) \cdot \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \mu_i(n_i + 1) \cdot 1_{\{k_i < b_i\}} \\
&+ \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j)\right) \cdot \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i \cdot 1_{\{k_i > 0\}}.
\end{aligned}$$

By substitution of (2.3.2) we obtain

$$\begin{aligned}
&\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k})\\
&\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\nu\cdot p_i\cdot 1_{\{k_i<b_i\}}\Big)\\
=&\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k})\cdot\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{k_i>0\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot\lambda_i\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\cdot\nu\cdot p_i\cdot 1_{\{k_i>0\}}.
\end{aligned}$$

Cancelling $\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$ and the sums with the terms $\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{k_i>0\}}$ on both sides of the equation leads to

$$\begin{aligned}
&\widetilde{\theta}\,(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\nu\cdot p_i\cdot 1_{\{k_i<b_i\}}\Big)\\
=&\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot\lambda_i\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\cdot\nu\cdot p_i\cdot 1_{\{k_i>0\}}.
\end{aligned}\tag{2.3.4}$$

The right-hand side of the last equation is

$$\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k})\cdot\nu\cdot p_i\cdot 1_{\{k_i<b_i\}}+\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k})\cdot\lambda_i\cdot 1_{\{k_i>0\}},$$

which is obviously the left-hand side.

Inspection of the system (2.3.4) reveals that it is a "generator equation", i.e. the global balance equation $\widetilde{\theta}\cdot\mathbf{Q}_{red}=\mathbf{0}$ for a suitably defined ergodic Markov process on state space K with "reduced generator" $\mathbf{Q}_{red}$. Because the Markov process generated by $\mathbf{Q}_{red}$ is irreducible the solution of (2.3.4) is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Remark 2.3.3. $\widetilde{\theta}\,(\mathbf{k})=\widetilde{\theta}(k_1,\ldots,k_J,k_{J+1})$ is obtained as a strictly positive solution of (2.3.4) which resembles the global balance equations of an artificial non-standard Gordon-Newell network of queues with $J+1$ nodes and $\sum_{j\in\overline{J}}b_j$ customers, exponentially distributed service times with rate λ_j for $k_j\le b_j$ and "∞" otherwise at node $j\in\{1,\ldots,J\}$ and with rate ν at node $J+1$ (cf. Figure 2.3.1). More precisely, it is a starlike network with $r(j,J+1)=1$, $j\in\overline{J}$, and branching probabilities $r(J+1,j)=p_j$, $j\in\overline{J}$.

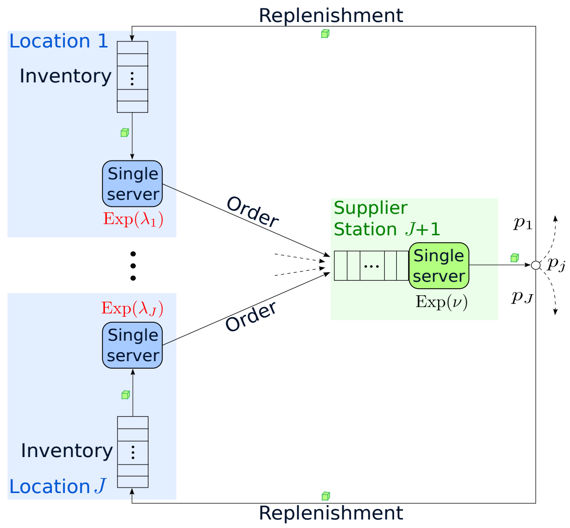


Figure 2.3.1.: Corresponding Gordon-Newell network

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 2.3.4. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E), \quad \pi(\mathbf{n},\mathbf{k}) := \lim_{t\to\infty} P(Z(t) = (\mathbf{n},\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := \left(\xi(\mathbf{n}) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi(\mathbf{n}) := \lim_{t\to\infty} P((X_1(t),\dots,X_J(t)) = \mathbf{n}),$$

$$\theta := (\theta(\mathbf{k}) : \mathbf{k} \in K), \quad \theta(\mathbf{k}) := \lim_{t\to\infty} P((Y_1(t),\dots,Y_J(t),W_{J+1}(t)) = \mathbf{k}).$$

Theorem 2.3.5. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n},\mathbf{k}) = \xi(\mathbf{n}) \cdot \theta(\mathbf{k}), \tag{2.3.5}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{2.3.6}$$

$$\theta(\mathbf{k}) = \theta(k_1,\dots,k_J,k_{J+1}) = C_\theta^{-1} \prod_{j\in\overline{J}} \left(\frac{\nu p_j}{\lambda_j}\right)^{k_j}, \quad \mathbf{k} \in K, \tag{2.3.7}$$

and normalisation constants

$$C_j = \sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)},\ j \in \overline{J}, \qquad \text{and} \qquad C_\theta = \sum_{\mathbf{k} \in K} \prod_{j \in \overline{J}} \left(\frac{\nu p_j}{\lambda_j}\right)^{k_j}.$$

Proof. Z is ergodic, if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ from Proposition 2.3.1 can be normalised (i.e. $\sum_{\mathbf{n} \in \mathbb{N}_0} \sum_{\mathbf{k} \in K} x(\mathbf{n}, \mathbf{k}) < \infty$). Because of Proposition 2.3.1 it holds

$$\sum_{\mathbf{n} \in \mathbb{N}_0} \sum_{\mathbf{k} \in K} x(\mathbf{n}, \mathbf{k}) = \sum_{\mathbf{n} \in \mathbb{N}_0} \widetilde{\xi}(\mathbf{n}) \cdot \sum_{\mathbf{k} \in K} \widetilde{\theta}(\mathbf{k}) = \left(\prod_{j \in \overline{J}} \sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}\right) \cdot \left(\sum_{k \in K} \prod_{j \in \overline{J}} \left(\frac{\nu p_j}{\lambda_j}\right)^{k_j}\right).$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 2.3.1 can be normalised if and only if $\sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty$ for all $j \in \overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n}, \mathbf{k}) = \frac{x(\mathbf{n}, \mathbf{k})}{\sum_{\mathbf{n} \in \mathbb{N}_0} \sum_{\mathbf{k} \in K} x(\mathbf{n}, \mathbf{k})},$$

where $x(\mathbf{n}, \mathbf{k})$ is given in Proposition 2.3.1. □

Remark 2.3.6. The expression (2.3.5) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

Representation (2.3.6) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(n) = \prod_{j \in \overline{J}} \xi_j(n_j) = \prod_{j \in \overline{J}} C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. In our production network, however, servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (2.3.5) has been unexpected to us.

Our production-inventory-replenishment system can be considered as a "Jackson network in a random environment" in [KDO16, Section 4]. We can interpret the inventory-replenishment subsystem, which contributes via θ to (2.3.5), as a "random environment" for the production network of nodes $\overline{J}$, which is a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1] we conclude from the hindsight that decoupling of the queueing process $(X_1, \ldots, X_J)$ and the process $(Y_1, \ldots, Y_J, W_{J+1})$, i.e. the formula (2.3.5), is a consequence of that Theorem 4.1.

Our direct proof of Theorem 2.3.5 is much shorter than embedding the present model into the general framework of [KDO16].

This page

intentionally

left blank.

2.4. Comparison with a more complex model with location specific items

In this section, we investigate the model of Section 2.2 as approximation for a more complex system where the central supplier sends raw materials to the locations that sequentially ordered them. This more complex system can be classified as a "multi-product system" because items are not exchangeable. Although we analyse this more complex model in Chapter 5 and derive an explicit solution for the stationary distribution in product form, the drawback of this multi-product model with deterministic routing is that the state space of the associated Markov process explodes. We therefore analyse the performance of the multi-product system with simulation and compare this to the results of the analytical expressions derived in Section 2.2 for the simpler model with exchangeable items and random routing.

Obviously, the main problem is to find correct values for the routing probabilities $(p_j : j = 1, \ldots, J)$. We have two approaches, first by simulation and second by an iterative algorithm. In both cases, we start from the fact that a portion $\Lambda_j := \lambda_{J+1} \cdot \check{p}_j$ is sent to location j from the overall departure rate λ_{J+1} (= throughput) of the central supplier for some $\check{p}_j \in (0, 1)$. In equilibrium this is exactly the replenishment rate originating at location j. The portion Λ_j can be obtained by simulations. The natural choice is then

$$p_j := \Lambda_j / (\sum_{k=1}^{J} \Lambda_k) = (\lambda_{J+1} \cdot \check{p}_j) / (\sum_{k=1}^{J} (\lambda_{J+1} \cdot \check{p}_k)) = \check{p}_j$$

because $\sum_{k=1}^{J} \check{p}_k = 1$.

(I) SIMULATION. Our first approach is to estimate $\check{p}_j$ using the transfer statistics from the central supplier to location j obtained in simulation runs of the multi-product model. Thereafter these estimated values are inserted into the analytical formulas of the simpler model. Clearly, this is not a practical recipe, but a way to get insights into possible similarities of both models. Great differences e.g. in inventory sizes or queue lengths would recommend not to use the simpler model. But fortunately enough, these differences are small.

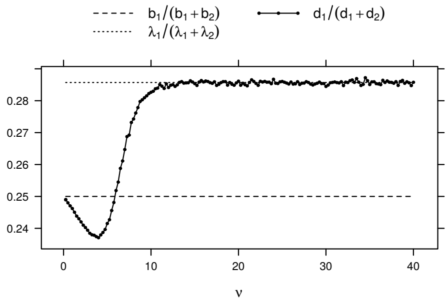
For this purpose we construct a fictional system with deterministic routing with two locations, $\overline{J} = \{1, 2\}$, and parameters $\lambda_1 = 2$, $\lambda_2 = 5$, $\mu_1 = 4$, $\mu_2 = 10$, $b_1 = 4$, $b_2 = 12$ and ν in a range from 0.5 to 40 with step size 0.25. We have chosen the service rates μ_j greater than the demand rates λ_j to keep the system ergodic, and we have chosen the demand rates λ_j and the base stock levels b_j in such a way that the ratios λ_1/λ_2 and b_1/b_2 are different. A larger base stock level b_2 is attributed to a system with a larger demand stream λ_2. Furthermore, we have chosen a very large run time $T = 100000$ to obtain the results close to the steady state solution in a single simulation run. We stored the number d_j of finished items sent from the central supplier to location j. The results of the simulation are plotted in Figure 2.4.1.

From Figure 2.4.1(a) we see, if ν is small, $d_1/(d_1 + d_2)$ seems to approach $b_1/(b_1 + b_2)$, and if ν is large, $d_1/(d_1 + d_2)$ approaches $\lambda_1/(\lambda_1 + \lambda_2)$. The intuitive explanation of this behaviour with the help of Figures 2.4.1(b) - 2.4.1(d) is as follows: If the central supplier is much slower than the production systems at the locations, then the central supplier's queue is almost always full and the inventories are almost always depleted. This means, in the queue of the central supplier there are approximately b_1 orders from location 1 and b_2 orders from location 2 in random order.

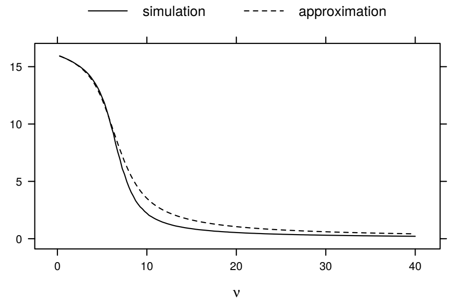
If the central supplier is much faster than the locations, then the orders pass the central supplier almost immediately. The order streams from locations 1 and 2 behave similarly to the superposition of two stochastically independent Poisson streams with rates λ_1 and λ_2. When two independent Poisson streams pass the central supplier with no delay, this is stochastically the same, as to input a Poisson stream with non-distinguishable orders to the central supplier with a rate $\lambda_1 + \lambda_2$ and then randomly decide of which type (location) it is: of type 1 with probability $\lambda_1/(\lambda_1 + \lambda_2)$ and of type 2 with probability $\lambda_2/(\lambda_1 + \lambda_2)$.

Figures 2.4.1(e) - 2.4.1(f) show that the analytically obtained average queue size almost perfectly matches the simulated (true) values of the average queue sizes at the locations. Furthermore, we can see that the average queue size is independent of the service rate ν at the central supplier (see equation (2.3.7)), which is predicted by the product form stationary distribution (2.3.5).

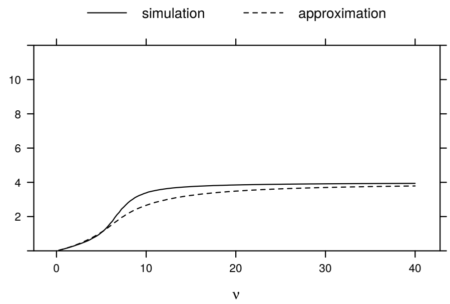
The comparison of the performance metrics of the simulated more complicated model with those of our analytically obtained results shows that the analytically obtained values can be used as an approximation for the multi-product and more complicated system's metrics. As noted before, the much more extensive comparisons of [GKAR12] support such substitutions of complex systems by suitably chosen product form systems as well.



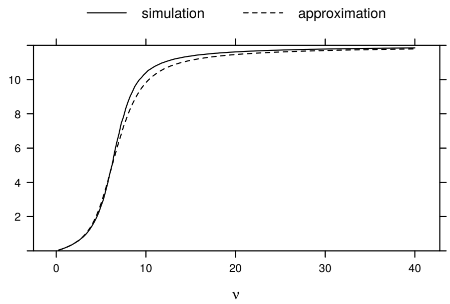
(a) $d_1/(d_1+d_2)$ — relative departure portion of finished items for location 1 from the central supplier in the system with deterministic routing



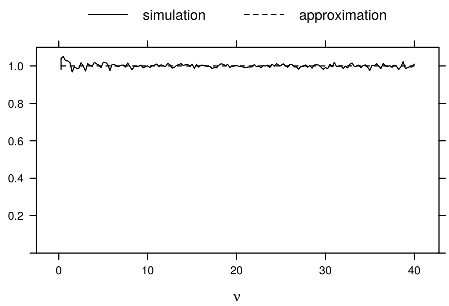
(b) Average queue size at the central supplier



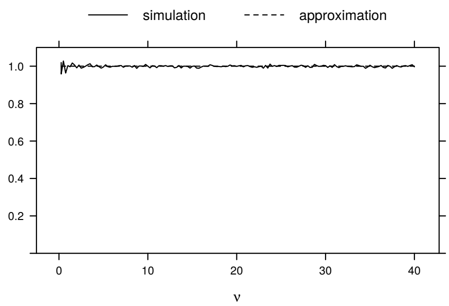
(c) Average inventory at location 1



(d) Average inventory at location 2



(e) Average queue size at location 1



(f) Average queue size at location 2

Figure 2.4.1.: Comparison between the simulated results of the system with deterministic routing and the analytic results of the system with random routing

(II) ITERATIVE ALGORITHM. In our second approach we use a queueing model to set the values of the routing probabilities p_j, $j \in \overline{J}$. In this procedure the routing probabilities are determined as being proportional to the effective arrival rates from location j at the central supplier. We consider the following model, where an admitted customer at location j represents an outstanding replenishment order from location j. The process time of such orders is complex (consisting of waiting and service times at the central supplier). Therefore, the admitted customer's service time is modelled as a random variable with a general distribution. Consequently, each location j, $j \in \overline{J}$, is considered as an $M/G/b_j/0$-Erlang-loss system[3] as depicted in Figure 2.4.2.

The arrival rate λ_j, $j = 1, \ldots, J$, is diminished by the loss probability q_j when all b_j units are on order at the central supplier. The value of q_j will be determined iteratively. Furthermore, the total arrival rate λ_{J+1} at the central supplier equals $\sum_{j=1}^{J} \lambda_j \cdot (1 - q_j)$.

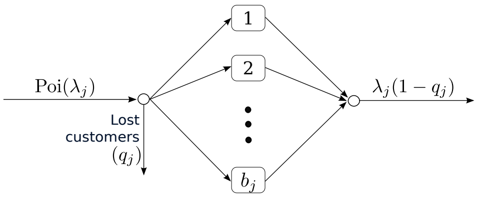


Figure 2.4.2.: Location j approximated as an $M/G/b_j/0$-FCFS queue

The central supplier is modelled as an $M/M/1/(b-1)$-FCFS queue with $b = \sum_{j \in \overline{J}} b_j$, service rate ν, and arrival rate λ_{J+1} of orders generated by admitted customers at the local $M/G/b_j/0$ queues (Figure 2.4.3).

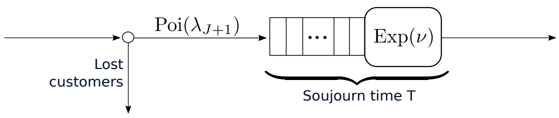


Figure 2.4.3.: Central supplier approximated as an $M/M/1/(b-1)$-FCFS queue

To determine the arrival rate at the central supplier, the blocking probabilities (loss probabilities) q_j of the $M/G/b_j/0$ queues need to be known and to determine these blocking probabilities we need to know the sojourn time T of the replenishment orders at the central supplier. This can be solved iteratively where the algorithm will stop if the blocking probabilities q_j remain unchanged under further iterations. As a result the routing probabilities p_j, $j \in \overline{J}$, can be calculated because they are proportional to the effective customer arrival rates at the central supplier. That is

$$p_j = \frac{(1 - q_j) \cdot \lambda_j}{\sum_{k \in \overline{J}} (1 - q_k) \cdot \lambda_k}.$$

[3] b_j service channels, no waiting room, Poisson arrivals, general service time distribution

---

Algorithm Calculation of p_j, $j \in \overline{J}$

Input:	number of locations	J
	service rate of the central supplier	ν
	arrival rates at the locations $j \in \overline{J}$	λ_j
	base stock levels at the locations $j \in \overline{J}$	b_j
	stop criterion	ε

Output: routing probabilities p_j, $j \in \overline{J}$

Initialize: blocking probabilities $q_j = 0, \ \forall j \in \overline{J}$

(1) Calculate the effective arrival rate of replenishment orders at the central supplier

$$\lambda_{J+1} = \sum_{j=0}^{J} (1 - q_j) \cdot \lambda_j$$

and the average sojourn time $E[T]$ of a replenishment order at the central supplier

$$E[T] = \frac{1}{\sum_{\ell=0}^{b} \left(\frac{\lambda_{J+1}}{\nu}\right)^{\ell}} \cdot \sum_{n=0}^{b-1} \left(\frac{\lambda_{J+1}}{\nu}\right)^{n} \cdot \frac{n+1}{\nu} \qquad \text{with } b = \sum_{j \in \overline{J}} b_j.$$

(2) Determine the new blocking probabilities $q_j^{(\text{new})}$ at the locations $j \in \overline{J}$

$$q_j^{(\text{new})} = \frac{(\lambda_j \cdot E[T])^{b_j}}{\sum_{n=0}^{b_j} (\lambda_j \cdot E[T])^n}.$$

(3) If

$$\sum_{j \in \overline{J}} |q_j - q_j^{(\text{new})}| > \varepsilon,$$

then

$$q_j \leftarrow q_j^{(\text{new})} \quad \forall j \in \overline{J} \qquad \text{and return to } \mathbf{(1)},$$

else calculate the routing probabilities p_j, $j \in \overline{J}$,

$$p_j = \frac{(1 - q_j^{(\text{new})}) \cdot \lambda_j}{\sum_{j \in \overline{J}} (1 - q_j^{(\text{new})}) \cdot \lambda_j}.$$

The iterative algorithm can be modelled in R. The R code is presented in Appendix B.1 on page 267.

We can use the simulation results of the multi-product system again and compare the values of p_j with the results of the iterative procedure. We again consider two locations $\overline{J} = \{1, 2\}$ with the same parameter values, except we do not need the service rates μ_j for the algorithm. We have chosen a stop criterion $\varepsilon = 0.001$. Furthermore, we stopped the algorithm when no convergence was reached within 500 iterations.

The results are plotted in Figure 2.4.4. We see that the resulting routing probabilities p_j of the algorithm show a good approximation if the service rate of the central supplier ν is larger than or equal to some value ν^*. In Figure 2.4.4 the value of ν^* is approximately 5.5. In the grey area of Figure 2.4.4, the iterative algorithm did not satisfy the stop criterion ε in less than 500 iterations for about 85% of the instances. The resulting values of p_j are not good approximations for the actual values. A series of experiments with λ_1 and λ_2 in a range from 0.5 to 20 with step size 0.25 supports our conjecture that such a value ν^* exists in general. The existence of ν^* is an open problem.

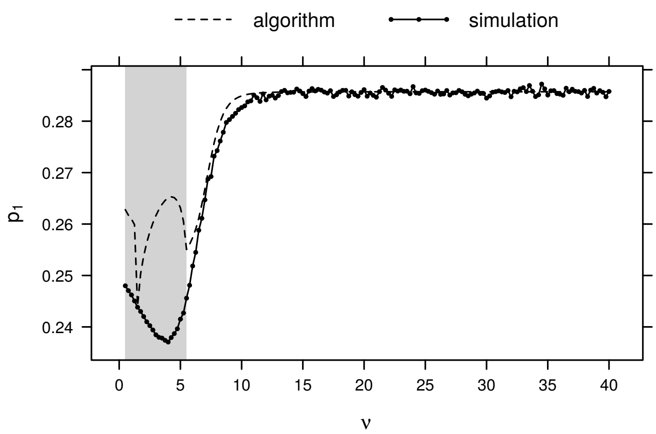


Figure 2.4.4.: Comparison between the simulated results and the results of the algorithm

We found it necessary to investigate the relationship between ν^* and other parameters of the system,[4] i.e. to determine the range of ν-values where the algorithm converges.

For $\lambda_1 = \lambda_2$ in a range from 0.5 to 20 with step size 0.25 (and all other parameters equal to the above) we see in Figure 2.4.5 that there is an approximately linear relationship between the arrival rates at the locations and ν^*.

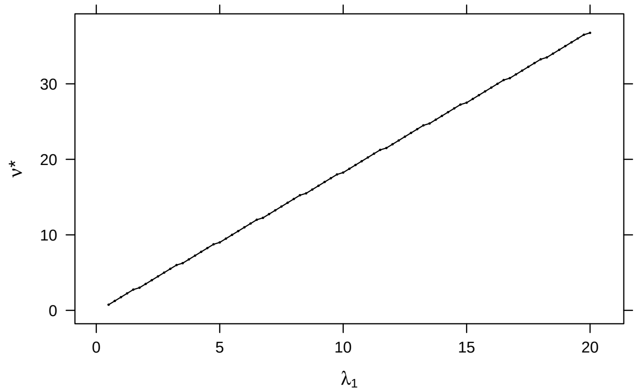
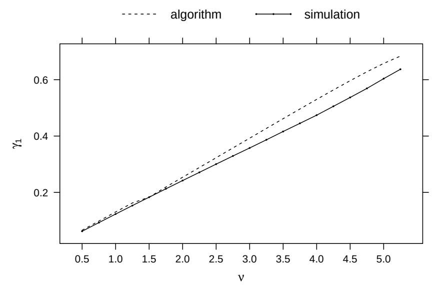


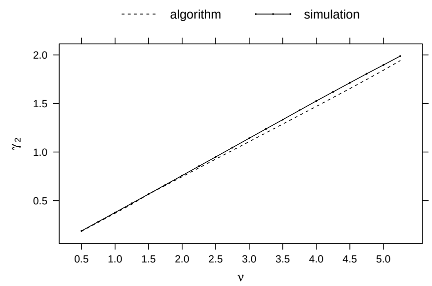
Figure 2.4.5.: Relationship between ν^* and $\lambda_1 = \lambda_2$

[4]The investigation of the relationship between ν^* and other parameters of the system is an improved version of [OKD16].

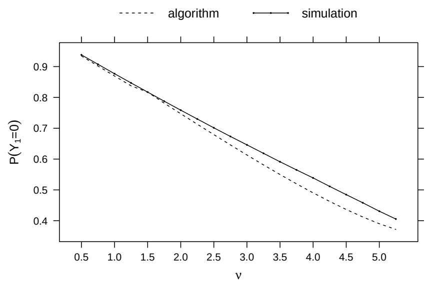
We mention that the optimal base stock levels b_j, which are the final decision variables, are determined in Theorem 2.5.2. The values p_j are needed in $\gamma_j = \frac{\nu p_j}{\lambda_j}$ to calculate $P(Y_j = 0)$ and $E(Y_j)$ in the cost function $\overline{g}_j(b_j)$. The decision in the grey area seems to be relatively robust because ν is small in this area. However, it depends on the combination of λ_j and ν. [5]We can use the simulation results of the multi-product system again and compare the values with the results of the iterative procedure with the routing probabilities p_j in the grey area where $\nu < \nu^*$. In our example the value of ν^* is approximately 5.5. The results are plotted in Figure 2.4.6.



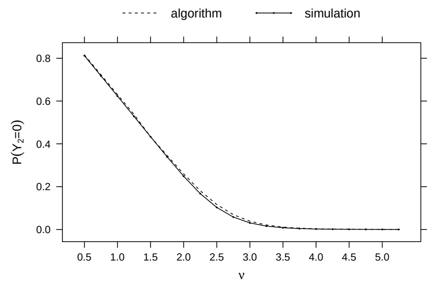
(a) $\gamma_1 = \frac{\nu p_1}{\lambda_1}$



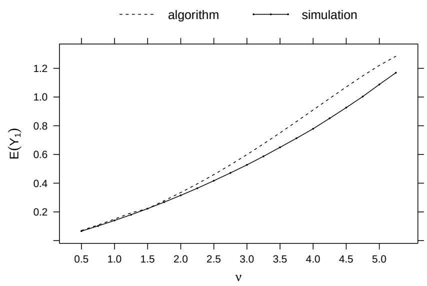
(b) $\gamma_2 = \frac{\nu p_2}{\lambda_2}$



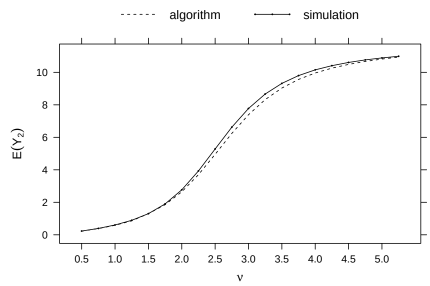
(c) $P(Y_1 = 0)$



(d) $P(Y_2 = 0)$



(e) $E(Y_1)$



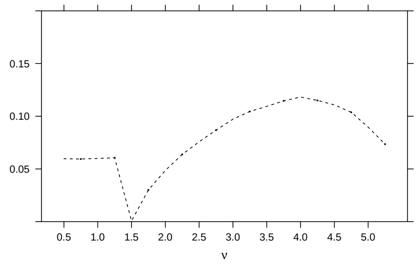
(f) $E(Y_2)$

Figure 2.4.6.: Comparison between the simulated results and the results of the algorithm in the grey area (i.e. ν in a range from 0.5 to 5.25 with step size 0.25)

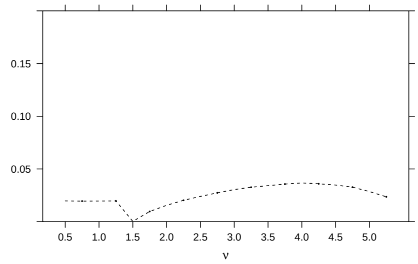
[5]The following investigations are an improved version of [OKD16].

Furthermore, we have plotted the relative errors $\frac{|\text{value of simulation}-\text{value of algorithm}|}{\text{value of simulation}}$ for γ_j, $P(Y_j = 0)$ and $E(Y_j)$ in Figure 2.4.7. We see the greatest relative error for $P(Y_2 = 0)$ in Figure 2.4.7(d) which was not clearly visible in Figure 2.4.6(d). However, it seems to be relatively robust since the value of $P(Y_2 = 0)$ is small in this area (see Figure 2.4.6(d)). Consequently, the resulting values for γ_j, $P(Y_j = 0)$ and $E(Y_j)$ show a good approximation for the actual values in our example.

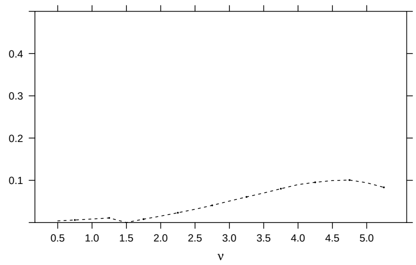
However, the decision of the optimal base stock levels also depends on the specific cost values. Therefore, further studies are still needed to make a statement about the robustness of the optimal base stock levels in the grey area.



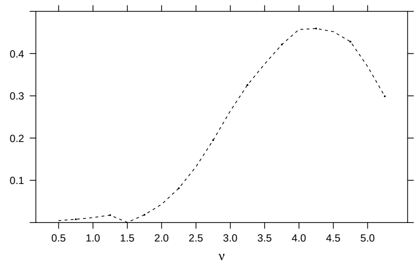
(a) $\gamma_1 = \frac{\nu p_1}{\lambda_1}$



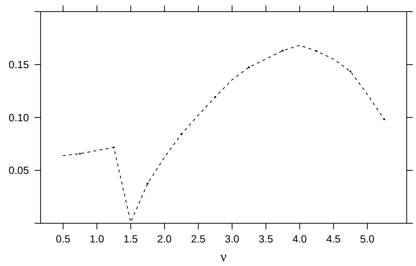
(b) $\gamma_2 = \frac{\nu p_2}{\lambda_2}$



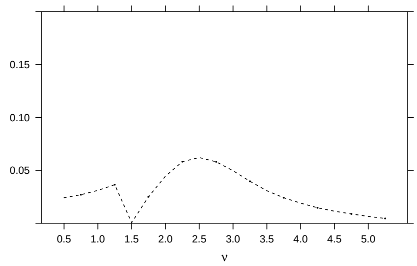
(c) $P(Y_1 = 0)$



(d) $P(Y_2 = 0)$



(e) $E(Y_1)$



(f) $E(Y_2)$

Figure 2.4.7.: Relative errors $\frac{|\text{value of simulation}-\text{value of algorithm}|}{\text{value of simulation}}$

2.5. Cost analysis

The total costs at location $j \in \overline{J}$ consist of shortage costs $c_{ls,j}$ for each customer that is lost, waiting costs $c_{w,j}$ per unit of time for each customer in the system (waiting or in service), capacity costs $c_{s,j}$ per unit of time for providing inventory storage space (e.g. rent, insurance), holding costs $c_{h,j}$ per unit of time for each unit that is kept on inventory. The unit holding costs at the central supplier are $c_{h,J+1}$. We assume that all of these costs per unit of time are positive.

Therefore, the cost function per unit of time in the respective states is[6]

$$f_{\mathbf{b}} : \mathbb{N}_0^J \times K \longrightarrow \mathbb{R}_0^+, \qquad f_{\mathbf{b}}(\mathbf{n}, \mathbf{k}) = \left(\sum_{j \in \bar{J}} f_{b_j}(n_j, k_j) + f_{J+1}(k_{J+1}) \right)$$

with the cost functions $f_{b_j} : \mathbb{N}_0^2 \longrightarrow \mathbb{R}_0^+$ at location j of the local system state (n_j, k_j) with base stock level b_j per unit of time

$$f_{b_j}(n_j, k_j) = c_{w,j} \cdot n_j + c_{s,j} \cdot b_j + c_{h,j} \cdot k_j + c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j = 0\}}$$

and the cost function $f_{J+1} : \mathbb{N}_0 \longrightarrow \mathbb{R}_0$ per unit of time at the central supplier

$$f_{J+1}(k_{J+1}) = c_{h,J+1} \cdot k_{J+1}.$$

We will analyse average long-term costs of the system as a function of the base stock levels $\mathbf{b} = (b_1, \ldots, b_J)$.

Lemma 2.5.1. *Optimal solutions for the problem described in Definition 2.2 are the set*

$$\arg\min\left(\overline{g}(\mathbf{b})\right)$$

with

$$\begin{aligned}
\overline{g}(\mathbf{b}) &:= \sum_{j \in \overline{J}} c_{s,j} \cdot b_j + \sum_{\mathbf{k} \in K} \left(\sum_{j \in \overline{J}} c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j = 0\}} + \sum_{j \in \overline{J} \cup \{J+1\}} c_{h,j} \cdot k_j \right) \cdot \theta(\mathbf{k}) \\
&= \sum_{j \in \overline{J}} (c_{s,j} + c_{h,J+1}) \cdot b_j + \sum_{\mathbf{k} \in K} \left(\sum_{j \in \overline{J}} \left(c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j = 0\}} + (c_{h,j} - c_{h,J+1}) \cdot k_j \right) \right) \cdot \theta(\mathbf{k}).
\end{aligned}$$

Proof. The asymptotic average costs for an ergodic system can be calculated as

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f_{\mathbf{b}}(Z(\omega, t)) dt = \sum_{(\mathbf{n}, \mathbf{k})} f_{\mathbf{b}}(\mathbf{n}, \mathbf{k}) \cdot \pi(\mathbf{n}, \mathbf{k}) =: \overline{f}(\mathbf{b}) \qquad P - a.s.$$

[6] The definition of the cost function in [OKD16] is corrected here.

Using product form properties of the system we obtain

$$
\begin{aligned}
&\overline{f}(\mathbf{b})\\
&= \sum_{(\mathbf{n},\mathbf{k})} \left(\sum_{j\in\overline{J}} f_{b_j}(n_j,k_j) + f_{J+1}(k_{J+1}) \right) \cdot \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right) \cdot \theta\left(\mathbf{k}\right)\\
&= \sum_{(\mathbf{n},\mathbf{k})} \left(\sum_{j\in\overline{J}} (c_{w,j}\cdot n_j + c_{s,j}\cdot b_j + c_{h,j}\cdot k_j + c_{ls,j}\cdot \lambda_j \cdot 1_{\{k_j=0\}}) + c_{h,J+1}\cdot k_{J+1} \right)\\
&\qquad \cdot \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right) \cdot \theta\Big(\mathbf{k}\Big)\\
&= \sum_{\mathbf{n}} \sum_{\mathbf{k}} \left(\sum_{j\in\overline{J}} (c_{w,j}\cdot n_j) \right) \cdot \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right) \cdot \theta(\mathbf{k})\\
&\quad + \sum_{\mathbf{n}} \sum_{\mathbf{k}} \left(\sum_{j\in\overline{J}} (c_{s,j}\cdot b_j + c_{h,j}\cdot k_j + c_{ls,j}\cdot \lambda_j \cdot 1_{\{k_j=0\}}) + c_{h,J+1}\cdot k_{J+1} \right)\\
&\qquad\quad \cdot \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right) \cdot \theta(\mathbf{k})\\
&= \underbrace{\sum_{\mathbf{k}} \theta(\mathbf{k})}_{=1} \cdot \sum_{\mathbf{n}} \sum_{j\in\overline{J}} c_{w,j}\cdot n_j \cdot \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right)\\
&\quad + \underbrace{\sum_{\mathbf{n}} \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right)}_{=1} \cdot \sum_{\mathbf{k}} \left(\sum_{j\in\overline{J}} (c_{s,j}\cdot b_j + c_{h,j}\cdot k_j + c_{ls,j}\cdot \lambda_j \cdot 1_{\{k_j=0\}}) + c_{h,J+1}\cdot k_{J+1} \right)\\
&\qquad\qquad \cdot \theta(\mathbf{k}).
\end{aligned}
$$

Let X_j, $j \in \overline{J}$, denote random variables which are distributed according to ξ_j. Using

$$
\begin{aligned}
\sum_{\mathbf{n}} \sum_{j\in\overline{J}} c_{w,j}\cdot n_j \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right) &= \sum_{j\in\overline{J}} \sum_{n_1=0}^{\infty} \cdots \sum_{n_j=0}^{\infty} \cdots \sum_{n_J=0}^{\infty} c_{w,j}\cdot n_j \cdot \left(\prod_{\ell\in\overline{J}} \xi_\ell(n_\ell) \right)\\
&= \sum_{j\in\overline{J}} \left(c_{w,j} \sum_{n_j=0}^{\infty} n_j \cdot \xi_j(n_j) \cdot \underbrace{\sum_{(n_i)_{i\neq j}} \left(\prod_{\ell\in\overline{J}\setminus\{j\}} \xi_\ell(n_\ell) \right)}_{=1} \right)\\
&= \sum_{j\in\overline{J}} c_{w,j} \cdot \sum_{n_j=0}^{\infty} n_j \cdot \xi_j(n_j)\\
&= \sum_{j\in\overline{J}} c_{w,j} \cdot E_{\xi_j}(X_j),
\end{aligned}
$$

we get for the asymptotic average costs

$$\overline{f}(\mathbf{b}) = \sum_{\mathbf{k}} \left(\sum_{j \in \overline{J}} (c_{s,j} \cdot b_j + c_{h,j} \cdot k_j + c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j=0\}}) + c_{h,J+1} \cdot k_{J+1} \right) \cdot \theta(\mathbf{k})$$
$$+ \underbrace{\sum_{j \in \overline{J}} c_{w,j} \cdot E_{\xi_j}(X_j)}_{\text{independent of } b_j}.$$

$$\Longrightarrow \arg\min \left(\overline{f}(\mathbf{b})\right) = \arg\min \left(\overline{g}(\mathbf{b})\right),$$

where

$$\begin{aligned}
\overline{g}(\mathbf{b}) &:= \sum_{\mathbf{k} \in K} \left(\sum_{j \in \overline{J}} (c_{s,j} \cdot b_j + c_{h,j} \cdot k_j + c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j=0\}}) + c_{h,J+1} \cdot k_{J+1} \right) \cdot \theta(\mathbf{k}) \\
&= \sum_{\mathbf{k} \in K} \left(\sum_{j \in \overline{J}} (c_{s,j} \cdot b_j + c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j=0\}}) + \sum_{j \in \overline{J} \cup \{J+1\}} c_{h,j} \cdot k_j \right) \cdot \theta(\mathbf{k}) \\
&= \sum_{j \in \overline{J}} c_{s,j} \cdot b_j + \sum_{\mathbf{k} \in K} \left(\sum_{j \in \overline{J}} c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j=0\}} + \sum_{j \in \overline{J} \cup \{J+1\}} c_{h,j} \cdot k_j \right) \cdot \theta(\mathbf{k}).
\end{aligned}$$

□

Although the locations and their describing processes are obviously strongly correlated because of the common replenishment mechanisms, it can be shown that this optimization problem is separable in the sense that we can split the global optimization problem into a set of independent local optimization problems.

Let $(Y_1, \dots, Y_J, Y_{J+1})$ denote random variables that are distributed according to θ. For improved readability we define

$$\gamma_j := \frac{\nu p_j}{\lambda_j}.$$

Theorem 2.5.2. *The optimal base stock levels* $\mathbf{b} = (b_1, \dots, b_J)$ *are determined as*

$$b_j \in \arg\min(\overline{g}_j) \quad \forall j \in \overline{J} \tag{2.5.1}$$

with

$$\overline{g}_j(b_j) := (c_{s,j} + c_{h,J+1}) \cdot b_j + c_{ls,j} \cdot \lambda_j \cdot P(Y_j = 0) + (c_{h,j} - c_{h,J+1}) \cdot E(Y_j),$$

where

$$P(Y_j = 0) = \left(\sum_{k_j=0}^{b_j} \gamma_j^{k_j} \right)^{-1} = \begin{cases} \frac{1-\gamma_j}{1-\gamma_j^{b_j+1}} & \text{for } \gamma_j \neq 1, \\ \frac{1}{b_j+1} & \text{for } \gamma_j = 1 \end{cases} \tag{2.5.2}$$

and

$$E(Y_j) = \left(\sum_{k_j=0}^{b_j} \gamma_j^{k_j} \right)^{-1} \sum_{k_j=0}^{b_j} k_j \cdot \gamma_j^{k_j} = \begin{cases} \frac{\gamma_j}{1-\gamma_j} \cdot \frac{b_j \cdot \gamma_j^{b_j+1} - (b_j+1) \cdot \gamma_j^{b_j} + 1}{1-\gamma_j^{b_j+1}} & \text{for } \gamma_j \neq 1, \\ \frac{b_j}{2} & \text{for } \gamma_j = 1. \end{cases} \tag{2.5.3}$$

Proof. We transform

$$\theta(\mathbf{k}) = \theta(k_1, \ldots, k_J, k_{J+1}) = C_\theta^{-1} \prod_{j \in \overline{J}} \gamma_j^{k_j}$$

on

$$K := \{(k_1, \ldots, k_J, k_{J+1}) | 0 \le k_j \le b_j,\ j = 1, \ldots, J,\ k_{J+1} = \sum_{j=1}^{J} (b_j - k_j)\} \subset \mathbb{N}_0^{J+1}$$

by an isomorphism to

$$\theta_-(\mathbf{k}_-) = \theta_-(k_1, \ldots, k_J) = C_\theta^{-1} \prod_{j=1}^{J} \gamma_j^{k_j} = \theta(\mathbf{k}) \tag{2.5.4}$$

on

$$K_- := \{(k_1, \ldots, k_J) | 0 \le k_j \le b_j,\ j \in \overline{J}\} = \prod_{j=1}^{J} \{0, 1, \ldots, b_j\}\,.$$

K_- is a product space and θ_- is a product measure. This leads to $\theta_-(\mathbf{k}_-) = \prod_{j=1}^{J} C_{\theta,j}^{-1} \cdot \gamma_j^{k_j}$ and

$$\begin{aligned}\overline{g}(\mathbf{b}) = &\sum_{j \in \overline{J}} (c_{s,j} + c_{h,J+1}) \cdot b_j \\ &+ \sum_{j \in \overline{J}} \left(\sum_{k_j=0}^{b_j} C_{\theta,j}^{-1} \cdot \gamma_j^{k_j} \cdot \left(c_{ls,j} \cdot \lambda_j \cdot 1_{\{k_j=0\}} + (c_{h,j} - c_{h,J+1}) \cdot k_j \right) \right. \\ &\qquad \left. \cdot \underbrace{\prod_{\substack{i=1 \\ i \ne j}}^{J} \sum_{k_i=0}^{b_i} C_{\theta,i}^{-1} \cdot \gamma_i^{k_i}}_{=1} \right).\end{aligned}$$

Set $(Y_1, \ldots, Y_J) \sim \theta_-$, then Y_j is distributed according to a truncated geometric distribution. It follows

$$\overline{g}(\mathbf{b}) = \sum_{j \in \overline{J}} \left((c_{s,j} + c_{h,J+1}) \cdot b_j + c_{ls,j} \cdot \lambda_j \cdot P(Y_j = 0) + (c_{h,j} - c_{h,J+1}) \cdot E(Y_j) \right).$$

□

We can show that a global optimal b_j^* exists.

Corollary 2.5.3. *For any $j \in \overline{J}$ the scaled costs $\frac{\overline{g}_j(b_j)}{b_j}$ are bounded above and below asymptotically*

$$0 < \liminf_{b_j \to \infty} \frac{\overline{g}_j(b_j)}{b_j} \le \limsup_{b_j \to \infty} \frac{\overline{g}_j(b_j)}{b_j} < \infty \tag{2.5.5}$$

and a global minimum b_j^* exists.

Proof. The property (2.5.5) follows from

$$\frac{\overline{g}_j(b_j)}{b_j} = \underbrace{c_{s,j}}_{>0} + c_{h,J+1} \cdot \left(1 - \underbrace{\frac{E(Y_j)}{b_j}}_{\in(0,1)}\right) + c_{ls,j} \cdot \lambda_j \cdot \underbrace{\frac{P(Y_j = 0)}{b_j}}_{\in(0,1)} + c_{h,j} \cdot \underbrace{\frac{E(Y_j)}{b_j}}_{\in(0,1)}.$$

The consequence of the last equation is $\lim_{b_j \to \infty} \overline{g}_j(b_j) = \infty$, which together with requirement $b_j \geq 1$ and $\overline{g}_j(b_j) > 0$ proves the existence of a global minimum. □

2.6. Structural properties of the integrated system

The investigations in this section rely on the fact that the product form of the stationary distribution (= separability) makes structures easily visible that are hard to detect by simulations or by direct numerical investigations. As a byproduct we demonstrate the power of product form calculus.

2.6.1. Ergodicity

As shown in Theorem 2.3.5, ergodicity is determined by $\sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty$, $j \in \overline{J}$, because K is finite. Hence, ergodicity is determined by the parameters of the isolated queueing system without the inventory system at the locations. For instance, if $\mu_j(\ell) = \mu_j, \forall \ell$, then $\lambda_j < \mu_j, \forall j$, is the correct condition for stabilizing the entire system.

Noteworthy is that the extra idle times of the servers at the production systems do not destroy ergodicity due to the necessary replenishments. The reason behind this is that the local lost sales at the individual servers at these locations automatically balance a possible bottleneck behaviour of the central supplier.

2.6.2. Effect of pooling demand, inventories and service capacity

Pooling of inventories and demand improves performance of inventory systems in many cases, for an early investigation see Eppen [Epp79]. Similarly, merging service capacities can improve throughput and reduce delay of production. Li and Zhang [LZ12] analyse how inventory pooling affects customer service levels. In our integrated production-inventory systems with a high degree of parallelism of the locations with demand, service, and inventory, it is therefore of interest whether pooling would improve the system's behaviour measured by the cost functions from Section 2.5. This is not a trivial question because for service systems (queues) van Dijk and van der Sluis [DS09] have shown that in many situations "Pooling is not the answer" to improve service quality. Similarly, for the classical newsvendor model there are situations where pooling of the demand is not an optimal decision. A review is given by Yang and Schrage [YS09] who show that pooling is not optimal in some cases. Bar-Lev and his coauthors [BL11] give precise necessary and sufficient conditions for the anomalies to occur or not to occur, i.e. pooling is suboptimal or optimal.

We utilize our closed-form formulas to show that pooling is favourable for some sets of parameter settings in the production-inventory system.

Definition 2.6.1. We say a subset $\{\phi_1, \ldots, \phi_N\} \subseteq \{1, \ldots, J\}$ of locations is homogeneous if

$$\lambda_{\phi_k} = \lambda_{\phi_\ell} \quad \text{and} \quad \gamma_{\phi_k} = \gamma_{\phi_\ell}, \quad \forall k, \ell \in \{1, \ldots, N\}.$$

A homogeneous set $\{\phi_1, \ldots, \phi_N\}$ of locations is 1-homogeneous if $\gamma_{\phi_k} = 1,\ \forall k \in \{1, \ldots, N\}$ holds.

Next we show that pooling of homogeneous systems (or locations) reduces the optimal base stock levels and the costs. This is done first for 1-homogeneous systems. We will verify this twofold. First by proving the following proposition.

Proposition 2.6.2. *Consider an ergodic network as in Section 2.2 that includes (among others) N locations with demand streams of intensities $\widehat{\lambda}_{\phi_1} = \cdots = \widehat{\lambda}_{\phi_N}$, with optimal base stock levels $\widehat{b}^*_{\phi_1}, \ldots, \widehat{b}^*_{\phi_N}$ and corresponding costs $\widehat{g}_{\phi_1}(\widehat{b}^*_{\phi_1}), \ldots, \widehat{g}_{\phi_N}(\widehat{b}^*_{\phi_N})$. Assume that these locations are 1-homogeneous, i.e. $\widehat{\gamma}_{\phi_i} = (\nu \widehat{p}_{\phi_i})/\widehat{\lambda}_{\phi_i} = 1,\ i = 1, \ldots, N$. When the N arrival streams are pooled to arrive at a single location, denoted by ϕ, with demand rate $\widehat{\lambda}_{\phi_1} + \cdots + \widehat{\lambda}_{\phi_N} = \widehat{\lambda}_{\phi_1} \cdot N =: \lambda_\phi$ and $p_\phi := \widehat{p}_{\phi_1} \cdot N$, then for the optimal base stock level b^*_ϕ and costs $\overline{g}_\phi(b^*_\phi)$ at the pooled location ϕ the following holds:*

$$\widehat{b}^*_{\phi_1} + \cdots + \widehat{b}^*_{\phi_N} \begin{cases} \approx \sqrt{N} \cdot b^*_\phi > b^*_\phi & \text{if } \widehat{b}^*_{\phi_1} > 1, \\ \gtrsim \sqrt{N} \cdot b^*_\phi > b^*_\phi & \text{if } \widehat{b}^*_{\phi_1} = 1 \end{cases}$$

and

$$\widehat{g}_{\phi_1}(\widehat{b}^*_{\phi_1}) + \cdots + \widehat{g}_{\phi_N}(\widehat{b}^*_{\phi_N}) \left\{ \begin{matrix} \approx \sqrt{N} \cdot \overline{g}_\phi(b^*_\phi) & \text{if } \widehat{b}^*_{\phi_1} > 1 \\ \gtrsim N \cdot \overline{g}_\phi(b^*_\phi) & \text{if } \widehat{b}^*_{\phi_1} = 1 \end{matrix} \right\} > \overline{g}_\phi(b^*_\phi).$$

An explanation for this decrease in case of pooling is probably that the system with more stations generates more variability in its performance metrics.

This proposition shows that the optimal base stock level and associated costs in the pooled system are smaller than the sum of the individual components in the non-pooled system.

Proof of Proposition 2.6.2. For technical reasons we investigate the reversed process of pooling: Splitting demand and inventory. For this we distribute the demand of rate λ_ϕ at location ϕ to N locations $\{\phi_1, \ldots, \phi_N\}$ with reduced demand of rate $\widehat{\lambda}_{\phi_k} := \lambda_\phi / N$ and $\widehat{p}_{\phi_k} := p_\phi / N,\ \forall k \in \{1, \ldots, N\}$.
We first collect necessary prerequisites and remark that $\gamma_\phi = 1$ implies $\widehat{\gamma}_{\phi_k} = 1,\ \forall k \in \{1, \ldots, N\}$ (and vice versa), i.e. the property of 1-homogeneity is hereditary for the N split locations.

For $\gamma_\phi = 1$ the cost function is

$$\begin{aligned} \overline{g}_\phi(b_\phi) &= (c_{s,\phi} + c_{h,J+1}) \cdot b_\phi + c_{ls,\phi} \cdot \lambda_\phi \cdot P(Y_\phi = 0) + (c_{h,\phi} - c_{h,J+1}) \cdot E(Y_\phi) \\ &\overset{(2.5.2)+(2.5.3)}{=} (c_{s,\phi} + c_{h,J+1}) \cdot b_\phi + c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{(b_\phi + 1)} + (c_{h,\phi} - c_{h,J+1}) \cdot \frac{b_\phi}{2} \\ &= \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1}) \right) \cdot b_\phi + c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{(b_\phi + 1)}. \end{aligned}$$

To simplify calculations we will analyse optimal points $b_\phi^* \in [1,\infty)$ in the continuous space. The first two derivatives of $\overline{g}_\phi(b_\phi)$ are

$$\frac{\partial \overline{g}_\phi}{\partial b_\phi}(b_\phi) = c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1}) - \frac{c_{ls,\phi} \cdot \lambda_\phi}{(b_\phi + 1)^2}, \quad \text{and} \quad \frac{\partial^2 \overline{g}_\phi}{\partial^2 b_\phi}(b_\phi) = \frac{2c_{ls,\phi} \cdot \lambda_\phi}{(b_\phi + 1)^3}.$$

Note that the second derivative is positive for $b_\phi \geq 1$. So the optimal base stock level b_ϕ^* for station ϕ with demand rate λ_ϕ is obtained from

$$\frac{\partial \overline{g}_\phi}{\partial b_\phi}(b_\phi^*) = 0 \wedge b_\phi^* > 1 \text{ or } b_\phi^* = 1 \quad \Longrightarrow \quad b_\phi^* = \max\left\{ \sqrt{\frac{c_{ls,\phi} \cdot \lambda_\phi}{c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})}} - 1,\ 1 \right\}.$$

According to Corollary 2.5.3 this single local minimum b_ϕ^* is also a global minimum.

To simplify notation we fix location ϕ_1 and compare the performance metrics and costs for arrival rate λ_ϕ and p_ϕ with the situation of reduced demand of rate $\widehat{\lambda}_{\phi_1} := \lambda_\phi / N$ and only a portion $\widehat{p}_{\phi_1} := p_\phi / N$ of items from the replenishment workstation being redirected to location ϕ_1. The quantities related to the location ϕ_1 will be tagged by a " $\widehat{\ }$ " and an index "$_{\phi_1}$". We have $\widehat{\gamma}_{\phi_1} = \gamma_\phi = 1$. All other cost values $\widehat{c}_{ls,\phi_1}$, $\widehat{c}_{s,\phi_1}$, $\widehat{c}_{h,\phi_1}$ remain the same as $c_{ls,\phi}$, $c_{s,\phi}$, $c_{h,\phi}$, and $c_{h,J+1}$ are already fixed. Using the previous results the new optimal base stock level $\widehat{b}_{\phi_1}^*$ for the location with reduced demand is

$$\widehat{b}_{\phi_1}^* = \max\left\{ \sqrt{\frac{c_{ls,\phi} \cdot \lambda_\phi / N}{c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})}} - 1,\ 1 \right\}. \tag{2.6.1}$$

Comparing the optimal base stock levels b_ϕ^* to $\widehat{b}_{\phi_1}^*$, we see that $b_\phi^* \geq \widehat{b}_{\phi_1}^*$. To be more precise,

$$\text{if } \widehat{b}_{\phi_1}^* > 1, \text{ then } \frac{\widehat{b}_{\phi_1}^* + 1}{b_\phi^* + 1} = \frac{1}{\sqrt{N}}, \quad \text{and if } \widehat{b}_{\phi_1}^* = 1, \text{ then } \frac{\widehat{b}_{\phi_1}^* + 1}{b_\phi^* + 1} > \frac{1}{\sqrt{N}}. \tag{2.6.2}$$

This implies for sufficiently large b_ϕ^*

$$\widehat{b}_{\phi_1}^* \approx \frac{b_\phi^*}{\sqrt{N}} > \frac{b_\phi^*}{N}, \text{ if } \widehat{b}_{\phi_1}^* > 1, \quad \text{and} \quad \widehat{b}_{\phi_1}^* \gtrsim \frac{b_\phi^*}{\sqrt{N}} > \frac{b_\phi^*}{N}, \text{ if } \widehat{b}_{\phi_1}^* = 1. \tag{2.6.3}$$

Equation (2.6.3) says that whenever the demand is scaled down by $1/N$ the optimal base stock level scales only with $1/\sqrt{N}$. This scaling is maintained for the standard costs which we consider (at least if $\widehat{b}_{\phi_1}^* > 1$). This follows from substituting $\widehat{b}_{\phi_1}^*$ into

$$\widehat{g}_{\phi_1}(\widehat{b}_{\phi_1}) := \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1}) \right) \cdot \widehat{b}_{\phi_1} + c_{ls,\phi} \cdot \widehat{\lambda}_{\phi_1} \cdot \frac{1}{(\widehat{b}_{\phi_1} + 1)}.$$

We directly obtain for $\widehat{b}_{\phi_1}^* > 1$

$$\widehat{g}_{\phi_1}(\widehat{b}_{\phi_1}^*) \approx \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1}) \right) \frac{b_\phi^*}{\sqrt{N}} + c_{ls,\phi} \cdot \frac{\lambda_\phi}{N} \cdot \frac{1}{\frac{(b_\phi^*+1)}{\sqrt{N}}} \approx \frac{1}{\sqrt{N}} \cdot \overline{g}_\phi(b_\phi^*) \tag{2.6.4}$$

and for $\widehat{b}^*_{\phi_1} = 1$

$$\widehat{g}_{\phi_1}(\widehat{b}^*_{\phi_1}) \gtrsim \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right) \frac{b^*_\phi}{N} + c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{b^*_\phi}{N} \gtrsim \frac{1}{N} \cdot \overline{g}_\phi(b^*_\phi). \tag{2.6.5}$$

□

Remark 2.6.3. $\approx$ resp. $\gtrsim$ in (2.6.3) means the following:

$$a(b^*_\phi) := \frac{\widehat{b}^*_{\phi_1} + 1}{b^*_\phi + 1} - \frac{\widehat{b}^*_{\phi_1}}{b^*_\phi} = \frac{\left(\widehat{b}^*_{\phi_1} + 1\right) \cdot b^*_\phi - \widehat{b}^*_{\phi_1} \cdot \left(b^*_\phi + 1\right)}{\left(b^*_\phi + 1\right) \cdot b^*_\phi} = \frac{b^*_\phi - \widehat{b}^*_{\phi_1}}{\left(b^*_\phi + 1\right) \cdot b^*_\phi} = \frac{1 - \frac{\widehat{b}^*_{\phi_1}}{b^*_\phi}}{b^*_\phi + 1}.$$

Because $b^*_\phi \geq \widehat{b}^*_{\phi_1}$ it follows that $\frac{\widehat{b}^*_{\phi_1}}{b^*_\phi} \leq 1$. This implies

$$\left| \frac{\widehat{b}^*_{\phi_1} + 1}{b^*_\phi + 1} - \frac{\widehat{b}^*_{\phi_1}}{b^*_\phi} \right| = \frac{\widehat{b}^*_{\phi_1} + 1}{b^*_\phi + 1} - \frac{\widehat{b}^*_{\phi_1}}{b^*_\phi} = \frac{1 - \frac{\widehat{b}^*_{\phi_1}}{b^*_\phi}}{b^*_\phi + 1} < \frac{1}{b^*_\phi + 1}.$$

Hence, for every $\varepsilon > 0$ we have for all $b^*_\phi \geq \max\left(1, \frac{1}{\varepsilon} - 1\right)$ that $0 < a(b^*_\phi) < \varepsilon$ which yields

$$\frac{\widehat{b}^*_{\phi_1}}{b^*_\phi} + a(b^*_\phi) = \frac{\widehat{b}^*_{\phi_1} + 1}{b^*_\phi + 1} \begin{cases} = \frac{1}{\sqrt{N}} & \text{if } \widehat{b}^*_{\phi_1} > 1, \\ > \frac{1}{\sqrt{N}} & \text{if } \widehat{b}^*_{\phi_1} = 1 \end{cases} \quad \Leftrightarrow \quad \widehat{b}^*_{\phi_1} \begin{cases} = \left(\frac{1}{\sqrt{N}} - a(b^*_\phi)\right) \cdot b^*_\phi & \text{if } \widehat{b}^*_{\phi_1} > 1, \\ > \left(\frac{1}{\sqrt{N}} - a(b^*_\phi)\right) \cdot b^*_\phi & \text{if } \widehat{b}^*_{\phi_1} = 1. \end{cases}$$

$\approx$ resp. $\gtrsim$ in (2.6.4) resp. (2.6.5) means the following:
Let $\varepsilon > 0$. Then we have for all $b^*_\phi \geq \max\left(1, \frac{1}{\varepsilon} - 1\right)$ that $0 < a(b^*_\phi) < \varepsilon$ and get for $\widehat{b}^*_{\phi_1} > 1$

$$\begin{aligned}
\widehat{g}_{\phi_1}(\widehat{b}^*_{\phi_1}) &= \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right) \left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \cdot b^*_\phi + c_{ls,\phi} \cdot \frac{\lambda_\phi}{N} \cdot \frac{1}{\frac{(b^*_\phi+1)}{\sqrt{N}}} \\
&= \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right) \left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \cdot b^*_\phi \\
&\quad + \left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1} + a\left(b^*_\phi\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1} \\
&= \left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \left(\left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right) \cdot b^*_\phi + c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1}\right) \\
&\quad + a\left(b^*_\phi\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1} \\
&= \left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \cdot \overline{g}_\phi(b^*_\phi) + a\left(b^*_\phi\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1}
\end{aligned}$$

and for $\widehat{b}^*_{\phi_1} = 1$

$$\begin{aligned}
\widehat{g}_{\phi_1}(\widehat{b}^*_{\phi_1}) &\geq \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right)\left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \cdot b^*_\phi + c_{ls,\phi} \cdot \frac{\lambda_\phi}{N} \cdot \frac{1}{\widehat{b}^*_{\phi_1} + 1} \\
&\overset{b^*_\phi \geq \widehat{b}^*_{\phi_1}}{\geq} \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right)\left(\frac{1}{\sqrt{N}} - a\left(b^*_\phi\right)\right) \cdot b^*_\phi + \frac{1}{N} \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1} \\
&\geq \left(c_{s,\phi} + \frac{1}{2}(c_{h,\phi} + c_{h,J+1})\right)\left(\frac{1}{N} - a\left(b^*_\phi\right)\right) \cdot b^*_\phi \\
&\quad + \left(\frac{1}{N} - a\left(b^*_\phi\right)\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1} + a\left(b^*_\phi\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1} \\
&= \left(\frac{1}{N} - a\left(b^*_\phi\right)\right) \cdot \overline{g}_\phi(b^*_\phi) + a\left(b^*_\phi\right) \cdot c_{ls,\phi} \cdot \lambda_\phi \cdot \frac{1}{b^*_\phi + 1}.
\end{aligned}$$

2.6.2.1. Pooling of general homogeneous locations: Refined numerical evaluation on the basis of product form structure

For the more general homogeneous system, we resort to a numerical investigation to verify whether equation (2.6.2) is still satisfied. To do so, we start with a system consisting of one location, then split it in two equal parts. Utilizing the separability and the decomposition property (2.5.1), we can reduce the problem to an isolated single location.

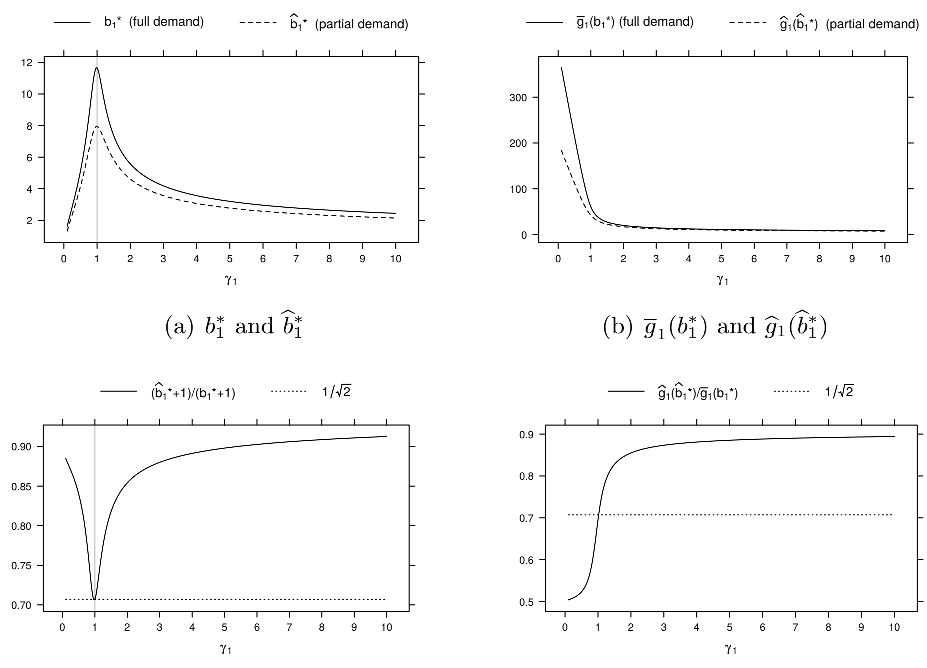
We consider the following fictional system with $\overline{J} = \{\phi\}$, $p_\phi = 1$, $\lambda_\phi = 1$, $c_{s,\phi} = 1$, $c_{h,\phi} = 2$, $c_{ls,\phi} = 400$, $c_{h,J+1} = 1$ and $\gamma_\phi \in [0.1, 10]$. The holding costs $c_{h,J+1}$ at the central supplier are lower than the holding costs $c_{h,\phi}$ at location ϕ. The very high cost $c_{ls,\phi} = 400$ results from expensive items, which justify the base stock policy, as argued in the introduction. We chose these numbers to obtain sufficiently large optimal base stock levels. When the location is split, such that the difference between continuous and discrete version of b^*_ϕ is negligible.

The results are plotted in Figure 2.6.1, where "full demand" refers to the original system (one location with $\lambda_\phi = 1$) and "partial demand" refers to one of the two split locations (which are identical, each with demand $\widehat{\lambda}_{\phi_1} = \widehat{\lambda}_{\phi_2} = 1/2$).

From Figure 2.6.1(a) we see that the values of b^*_ϕ and $\widehat{b}^*_{\phi_1}$ are highest, when $\gamma_\phi = 1$. Figure 2.6.1(b) demonstrates monotone decreasing behaviour of the cost functions $\overline{g}_\phi$, and $\widehat{g}_{\phi_1}$ in γ_ϕ. Figure 2.6.1(c) shows that $\frac{\widehat{b}^*_{\phi_1}+1}{b^*_\phi+1} \geq \frac{1}{\sqrt{2}}$. The ratio is close to its lowest value when $\gamma_\phi = 1$. That means that if γ_ϕ is only slightly different from 1, we soon gain more than factor $\frac{1}{\sqrt{2}}$ when two locations are pooled.

If γ_ϕ deviates more from 1 we observe that $\frac{\widehat{b}^*_{\phi_1}+1}{b^*_\phi+1} \approx 0.9$. The consequence is that pooling two demand streams of equal rate $\frac{\lambda_\phi}{2}$ yields a reduction of the needed inventory by a factor close to $1/2$.

This observation and conclusion is supported further by Figure 2.6.1(a). It is shown that for high replenishment rate $\nu = \frac{\gamma_\phi \cdot \lambda_\phi}{p_\phi}$ (with $\lambda_\phi = 1$ fixed), the optimal base stock level at demand rate λ_ϕ and $\frac{\lambda_\phi}{2}$ are almost the same. Finally, from Figure 2.6.1(d) we conclude that pooling two identical locations in the homogeneous case for large γ_ϕ reduced the optimal total costs by a factor $1/2$, too.



(a) b_1^* and $\widehat{b}_1^*$

(b) $\overline{g}_1(b_1^*)$ and $\widehat{g}_1(\widehat{b}_1^*)$

(c) $\frac{\widehat{b}_1^*+1}{b_1^*+1}$

(d) $\frac{\widehat{g}_1(\widehat{b}_1^*)}{\overline{g}_1(b_1^*)}$

Figure 2.6.1.: The optimal base stock levels of the original system b_1^* and of the splitted systems $\widehat{b}_1^*$ with corresponding optimal costs $\overline{g}_1(b_1^*)$ and $\widehat{g}_1(\widehat{b}_1^*)$ from Section 2.6.2.

Summarizing: From pooling two homogeneous locations we can expect roughly *at least* a gain of inventory reduction by a factor of $1/\sqrt{2}$, which is attained (approximately) in the 1-homogeneous case. Hence, for a subset of the parameter space, pooling is advisable, i.e. the "type-F-anomaly" [YS09, Section 2] does not occur.

Additional comments:

(1) The explicit product form expressions for the stationary distribution allow a more refined evaluation. We report only some interesting observations, which refer to an evaluation in the continuous optimization domain.

  (i) Although Figure 2.6.1(a) suggests that the optimal base stock levels b_ϕ^* and $\widehat{b}_{\phi_1}^*$ are maximal at $\gamma_\phi = 1$, we want to stress that this is in general *not* the case.

  (ii) Although Figure 2.6.1(c) suggests that the quotient $\frac{\widehat{b}_{\phi_1}^*+1}{b_\phi^*+1}$ is minimal at $\gamma_\phi = 1$, this is in general *not* the case as well.

  (iii) Figure 2.6.1(c) suggests even more that $\frac{1}{\sqrt{2}}$ is always a lower bound for $\frac{\widehat{b}_{\phi_1}^*+1}{b_\phi^*+1}$. We performed a detailed numerical evaluation with parameters $\lambda_\phi = 1$, $p_\phi = 1$, $c_{s,\phi}$, $c_{h,\phi}$, $c_{h,J+1}$ and ν from $\{0.1, 0.2, \ldots, 0.9\} \cup \{1, 2, \ldots, 9\} \cup \{10, 20, \ldots, 100\}$ and $c_{ls,\phi}$ from $\{0, 10, 20, \ldots, 1000\}$, which resulted in $62,080,256$ different system settings. They showed that the quotient $\frac{\widehat{b}^*_{\phi_1}+1}{b^*_\phi+1}$ fell below the value $\frac{0.99}{\sqrt{2}}$ for less than 6% of the instances, and below $\frac{0.95}{\sqrt{2}}$ for less than 0.5% of the instances. The smallest quotient was approximately $\frac{0.86}{\sqrt{2}}$.

(2) The strong peaks in Figure 2.6.1 (near $\gamma_\phi = 1$) are still waiting for an intuitive explanation. We emphasize that the possibility to detect these peaks strongly relied on the fact that due to the product form stationary distribution we could separate parts of the system (queues at locations) from other parts (inventory and central supplier).

2.6.3. Transformation of the stationary distribution

We started with evaluation of $\pi(\mathbf{n},\mathbf{k}) = \left(\prod_{j\in\overline{J}} \xi_j(n_j)\right) \cdot \theta(\mathbf{k})$, which made it easy to define and understand the cost structure of the system. Introducing later on for θ the isomorphic θ_- (see (2.5.4) on page 32) offers additional valuable insight into the structure of the inventory-supplier part of the integrated system.

Consider a situation where the service times for production of raw material at the workstation of the central supplier are extremely long (by chance). Then it is intuitive that the on-hand inventory levels at all locations are low (and base stock levels are high to prevent stockouts). Similarly, for short replenishment production times we see high stock levels at all locations. Therefore, it is a tempting conjecture that the inventory levels are *positively correlated*.

θ_- tells us that this intuition goes wrong in the long-run and in equilibrium: Inventory levels are *independent for any fixed time point*. Clearly, the inventory processes are *not* independent over time.

2.6.4. Monotonicity properties

[RW96] considered systems similar to ours (without service at the locations) and found by numerical studies that the optimal base stock levels in their network depend on the utilization of the replenishment server in a monotone fashion. From our product form equilibrium, monotonicity properties of various quantities can be derived analytically exploiting the form of $\pi(\mathbf{n},\mathbf{k})$ with the aid of stochastic order theory. This is a classical approach and seems to go back in inventory theory to the nineteen-sixties [Kar60]. We sketch two prototype examples:

(1) Recall that $(Y_1,\ldots,Y_J)$ is a vector distributed like θ_-. An intuitive conjecture is that: Increasing ν, decreasing λ_j or increasing p_j will increase $E(Y_j)$ and decrease $P(Y_j = 0)$, which occur in the local cost function $\overline{g}_j$.

This can be seen by proving that increasing γ_j to $\tilde{\gamma}_j$, which follows from either of the mentioned changes, implies stochastic ordering $Y_j \leq_{st} \tilde{Y}_j$ (with self-explaining notation). This is not directly visible due to the normalisation constants, but it is immediate that the likelihood ratio ordering $Y_j \leq_{lr} \tilde{Y}_j$ holds, which implies stochastic ordering. This leads to $E(Y_j) \leq E(\tilde{Y}_j)$ and $P(Y_j = 0) \geq P(\tilde{Y}_j = 0)$. See [MS02, p. 12] for details.

In a similar way the vector $(Y_1, \ldots, Y_J)$ of the joint inventory sizes with parameters γ_j is dominated in the sense of multivariate strong likelihood order ($= tp_2$ order) by a vector $(\tilde{Y}_1, \ldots, \tilde{Y}_J)$ of stock sizes in a system with parameters $\tilde{\gamma}_j \geq \gamma_j$, which implies multivariate stochastic ordering of the vectors $(Y_1, \ldots, Y_J) \leq_{st} (\tilde{Y}_1, \ldots, \tilde{Y}_J)$, see [MS02, pp. 129f.] for details.

A consequence of this observation is that whenever the demand intensity increases from λ_j to $\tilde{\lambda}_j$ with $\tilde{\lambda}_j \geq \lambda_j$ at some location j, the inventory position at the other locations will not decrease in the sense of (multivariate) stochastic order, which again has implications on the local cost functions via $E(Y_i)$ and decreases $P(Y_i = 0)$, as described above.

(2) A similar monotonicity prevails for the joint queue lengths vector $(X_1, \ldots, X_J)$, which in steady state behaves at fixed time instants like a vector of independent birth-death processes. The reasoning is the same as in (1).

2.6.5. Insensitivity and robustness

Sensitivity analysis is an important topic in classical inventory theory and is often hard to perform. In our model the stationary distribution $\pi(\mathbf{n}, \mathbf{k}) = \left(\prod_{j \in \overline{J}} \xi_j(n_j)\right) \cdot \theta(\mathbf{k})$ reveals strong insensitivity properties of the system which make sensitivity analysis amenable: The steady state behaviour of the subnetwork consisting of inventories and the central supplier does not change when the service rates at the locations are changed as long as the global system remains ergodic. Therefore θ is robust against estimation errors in determining the $\mu_j(\cdot)$. Vice versa, the distribution $\xi_j(n_j)$ is robust against changes in the inventory-supplier network as long as the demand intensity and the service rates are maintained.

3. Load balancing policies

In this chapter, we analyse the basic model, which we have introduced in Chapter 2, where routing of items depends on the on-hand inventory at the locations (with the aim to obtain "load balancing"). The systems under investigation differ with respect to the load balancing policy: In Section 3.3, we consider strict priorities (i.e. the finished item is sent to the location(s) with the highest difference between the on-hand inventory and the capacity of the inventory) and in Section 3.4, we consider weak priorities (i.e. the finished item is sent with greater probability to the location with higher difference between the on-hand inventory and the capacity of the inventory).

3.1. Related literature and own contributions

The research of such systems is motivated by state-dependent routing/branching of customers. The "(...) purpose of introducing flexible state-dependent routing strategies is to optimally utilize network resources and to minimise network delay and response times" [Dad87, p. 1] and "(...) that introducing state-dependent routing into product form networks usually destroys the product-form of the steady state probabilities" [Dad87, p. 1]. Product form solutions under state-dependent routing are, for example, found in [Pit79], [HD84], [Sch84].

Several other stock allocation policies can, for example, be found in the article of Abouee-Mehrizi and his coauthors [AMBB14]. They consider a two-echelon inventory system with a capacitated centralized production facility and several distribution centres. We will not go into any greater detail in this allocation policies.

Daduna [Dad85, p. 624] and Towsley [Tow80, pp. 327f.] argued that an optimal routing/branching policy for systems with identical peripheral processors is by intuitive reasoning: "Customers enter the peripheral processor with the shortest queue". This is equivalent to our strict priorities for load balancing policy.

Literature about strict priorities for load balancing policy:
Chow and Kohler analyse the performance of two-processor distributed computer systems under several dynamic load balancing policies in [CK77] and [CK79]. They compare the performance and their results indicate that a simple load balancing policy can significantly improve the performance (turnaround time) of the system.

In [CK77] they analyse the performance of homogeneous (i.e. identical) two-processor distributed computer systems under several dynamic load balancing policies. Their analysis is based on the recursive solution technique and they illustrated the algorithm for a sample system [CK77, Appendix, pp. 51f.]. Their strategy "join the shorter queue without channel transfer" in Chow and Kohler's Model B [CK77, pp. 42f.] is equivalent to our strict load balancing policy.

They mention that "the recursive solution technique can be applied to queueing models with more than one queue" [CK77, p. 45]. Furthermore, they "(...) are currently working to extend the analysis to (...) systems with more than two processors" [CK77, p. 49].[1]

Flatto and McKean study also Chow and Kohler's Model B in [FM77] and derive by the generating function approach a complicated closed form solution (cf. [FM77, Section 3, p. 261]), where the inter-arrival times of customers are exponentially distributed with rate 1.

Chow and Kohler present in [CK79] a generalization of the recursive solution technique. They apply the method to non-homogeneous (= heterogeneous) two-processor systems with special properties in [CK79, Section IV, pp. 358f.] and present a sample system using the algorithm in [CK79, Appendix, pp. 360f.]. They mention that the generalization of recursive solution "(...) technique for three or more processors does not appear to be straightforward" [CK79, p. 359].

Literature about weak priorities for load balancing policy:
Towsley analyses in [Tow75, Example 4.4.3, pp. 63ff. and Appendix D, pp. 138ff.] and in [Tow80, Example 1, pp. 327ff.] a central server model (called starlike system) with two identical processors, N customers, exponential service times, FCFS discipline and the following routing strategy (shortest queue = strict priority): A job leaving the central server goes to the processor with the shortest queue length. If both have the same queue length, it enters either with equal probability. He mentions that a numerical solution can be obtained by solving a discrete-state continuous-transition Markov model.

Towsley compares the performance of this model under different routing policies: constant and functional branching. His investigations show for a system with two identical processors that the use of functional branching probabilities is an excellent approximation for the "shortest queue"-system. The functional branching probabilities coincides with our load balancing policy with weak priorities and the "join the shortest queue"-policy is similar to our load balancing policy with strict priorities in our inventory-replenishment subsystem. His results support that functional branching probabilities can be used "(...) in modelling real-world problems such as load balancing" [Tow80, p. 328].

[1] We have found a reference to the article

> *Models for dynamic load balancing in homogeneous multiple processing systems* by Y.C. Chow and W.H. Kohler (IEE Transactions on Computers, volume c–36, pages 667–679)

in three sources:

> *Load Balancing in Parallel Computers: Theory and Practice* by C. Xu, F.C.M. Lau (Springer Science & Business Media, 1996)

> *Adaptive Load Sharing in Heterogenous Systems: Policies, Modifications, and Simulation* by K.Y. Kabalan, W.W. Smari and J.Y. Hakimian (International Journal of Simulation, Systems, Science and Technology, volume 3, number 1–2, pages 9–100, 2002)

> *Achieving High Performance on Extremely Large Parallel Machines: Performance Prediction and Load Balancing* by G. Zheng (Dissertation, University of Illinois at Urbana-Champaign, 2005)

When we were searching for this article in IEEE Xplore Digital Library, there was no article with this title. We can only find the article about heterogeneous multiple processing systems. We contacted the author Professor Yuan Chow and additionally Dr. Zheng, who has referred this article in his dissertation. But we obtained no precise information.

Towsley finds in [Tow80] for closed queueing networks a class of network topologies and a versatile class of state-dependent routing probabilities (branching probabilities) which lead to a product form equilibrium distribution. Further investigations on product form distribution and cycle time can be found in [Dad85] and [Dad87].

Our main contributions are the following:
For the system with *strict priorities* for load balancing policy we develop a Markov process. We prove that the stationary distribution has a product form of the marginal distributions of the production subsystem and of the inventory-replenishment subsystem.

We derive an explicit solution for the marginal distribution of the production subsystem. Furthermore, for some special cases we derive an explicit solution for the marginal distribution of the inventory-replenishment subsystem: For the case of J homogeneous locations (i.e. equal arrival rates and any service rates) with base stock levels equal to one, as well as for the case of two heterogeneous locations (i.e. any arrival and service rates) with base stock levels equal to one. For systems with base stock levels greater than one the marginal distribution of the inventory-replenishment subsystem with two locations can be obtained by a recursive method which is described by an algorithm.

Our work is an extension of the investigations of Chow and Kohler [CK77, CK79]: Their study is limited to two processors (= our inventories without production systems).

For the heterogeneous case, our load balancing policy is slightly different from that of Chow and Kohler [CK79, Section IV, pp. 358f.] (for more details see Appendix C.1 on page 272). Therefore, we construct a new algorithm.

For the system with *weak priorities* for load balancing policy we develop a Markov process and derive the stationary distribution in explicit product form. Our work is an extension of the investigations of Towsley [Tow75, Tow80] and Daduna [Dad85, Dad87]: Their closed network with branching policies is equivalent to our inventory-replenishment subsystem, which we have integrated in a complex supply chain with a production subsystem.

A cost analysis can be performed as for the basic model in Section 2.5 on page 29.

3.2. Description of the general model

The supply chain of interest is depicted in Figure 3.2.1. We have a set of locations $\overline{J} := \{1, 2, \ldots, J\}$, $J > 1$. Each of the locations consists of a production system with an attached inventory. The inventories are replenished by a single central supplier, which is referred to as workstation $J + 1$ and manufactures raw material for all locations. The items of raw material are indistinguishable (exchangeable).

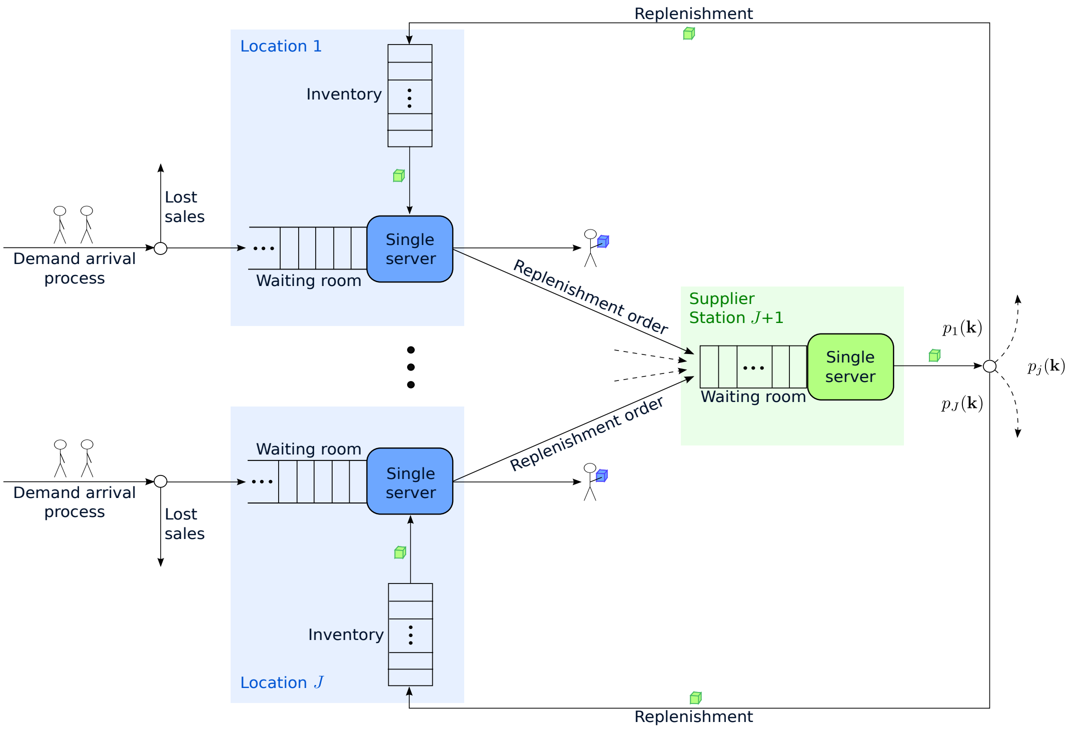


Figure 3.2.1.: Supply chain with base stock policies and load balancing policies

Facilities in the supply chain. Each production system $j \in \overline{J}$ consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under a FCFS regime. Customers arrive one by one at the production system j according to a Poisson process with rate $\lambda_j > 0$ and require service. To satisfy a customer's demand the production system needs exactly one item of raw material, which is taken from the associated local inventory. When a new customer arrives at a location while the previous customers' order is not finished, this customer will wait. If the inventory is depleted at location j, the customers who are already waiting in line will wait, but new arriving customers at this location will decide not to join the queue and are lost ("local lost sales").

The service requests at the locations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at location $j \in \overline{J}$ is provided with local queue-length-dependent intensity. If there are $n_j > 0$ customers present at location j, either waiting or in service (if any), and if the inventory is not depleted, the service intensity is $\mu_j(n_j) > 0$.

If the server is ready to serve a customer who is at the head of the line, and the inventory is not depleted, the service immediately starts. Otherwise, the service starts at the instant of time when the next replenishment arrives at the local inventory.

The inventory at location $j \in \overline{J}$ is controlled by prescribing a local base stock level $b_j \geq 1$, which is the maximal size of the inventory there, we denote $\mathbf{b} := \big(b_j : j \in \overline{J}\big)$.

The central supplier (which is referred to as workstation $J+1$) consists of a single server (machine) and a waiting room under FCFS regime. At most $\sum_{j\in\overline{J}} b_j - 1$ replenishment orders are waiting at the central supplier. Service times at the central supplier are exponentially distributed with parameter $\nu > 0$.

Routing in the supply chain. A served customer departs from the system immediately after service and the associated consumed raw material is removed from the inventory and an order of one item is placed at the central supplier at this time instant ("base stock policy").

A finished item of raw material departs from the central supplier immediately and is sent with probability $p_j(\mathbf{k})$, independent of the network's history, given $\mathbf{k}$, to location j, $j \in \overline{J}$, if the state of the inventory-replenishment subsystem is $\mathbf{k}$.
The systems under investigation differ with respect to the *load balancing policy* in the following way:

- Strict priorities (Section 3.3):
  The finished item of raw material is sent to the location(s) with the highest difference between the on-hand inventory and the capacity of the inventory (= base stock level).
- Weak priorities (Section 3.4):
  The finished item of raw material is sent with greater probability to the location with higher difference between the on-hand inventory and the capacity of the inventory (= base stock level).

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the central supplier and the local inventories are negligible.

The usual independence assumptions are assumed to hold as well.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$, either waiting or in service (queue length). By $Y_j(t)$ we denote the size of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_{J+1}(t)$ we denote the number of replenishment orders at the central supplier at time $t \geq 0$, either waiting or in service (queue length).

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), \ldots, X_J(t), Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) \,:\, t \geq 0)\,.$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular.

The state space of Z is

$$E = \left\{(\mathbf{n},\mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J,\ \mathbf{k} \in K\right\}$$

with

$$K := \left\{(k_1,\dots,k_J,k_{J+1}) | 0 \le k_j \le b_j,\ j=1,\dots,J,\ k_{J+1} = \sum_{j=1}^{J}(b_j - k_j)\right\} \subset \mathbb{N}_0^{J+1}.$$

Note the redundancy in the state space: $W_{J+1}(t) = \sum_{j\in\overline{J}} b_j - \sum_{j\in\overline{J}} Y_j(t)$. We prefer to carry all information explicitly with because the dynamics of the system are easier visible.

Definition 3.2.1. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E), \quad \pi(\mathbf{n},\mathbf{k}) := \lim_{t\to\infty} P(Z(t) = (\mathbf{n},\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := \left(\xi(\mathbf{n}) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi(\mathbf{n}) := \lim_{t\to\infty} P((X_1(t),\dots,X_J(t)) = \mathbf{n}),$$

$$\theta := (\theta(\mathbf{k}) : \mathbf{k} \in K), \quad \theta(\mathbf{k}) := \lim_{t\to\infty} P((Y_1(t),\dots,Y_J(t),W_{J+1}(t)) = \mathbf{k}).$$

3.3. Load balancing policy: Strict priorities

In this section, we study the supply chain with strict priorities as load balancing policy as described in Section 3.2.

We define

$$\begin{aligned}\arg\max_{i\in\overline{J}}(b_i - k_i) :&= \left\{i \in \overline{J} | (b_i - k_i) \text{ maximal}\right\} \\ &= \left\{i \in \overline{J} | \forall j \in \overline{J} :\ (b_i - k_i) \ge (b_j - k_j)\right\}.\end{aligned}$$

The finished item of raw material is sent to location $i \in \overline{J}$ with probability

$$p_i(\mathbf{k}) = \begin{cases} 1, & \text{if } \{i\} = \arg\max_{j\in\overline{J}}(b_j - k_j), \\ \frac{1}{|\arg\max_{j\in\overline{J}}(b_j-k_j)|} < 1, & \text{if } \{i\} \subsetneq \arg\max_{j\in\overline{J}}(b_j - k_j), \\ 0, & \text{if } i \notin \arg\max_{j\in\overline{J}}(b_j - k_j), \end{cases}$$

i.e. to the location(s) with the highest difference between the on-hand inventory and the capacity of the inventory (= base stock level), if the inventory is not full at this/these location(s) (this means that the on-hand inventory level at this/these location(s) is lower than the base stock level). The routing probabilities out of the central supplier must sum to one if there is at least one order at the central supplier.

3.3.1. Limiting and stationary distribution

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n},\mathbf{k}) \in E$:

$$q\left((\mathbf{n},\mathbf{k});(\mathbf{n}+\mathbf{e}_i,\mathbf{k})\right) = \lambda_i \cdot 1_{\{k_i>0\}}, \qquad i \in \overline{J},$$
$$q\left((\mathbf{n},\mathbf{k});(\mathbf{n}-\mathbf{e}_i,\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\right) = \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{k_i>0\}}, \qquad i \in \overline{J},$$
$$q\left((\mathbf{n},\mathbf{k});(\mathbf{n},\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\right) = \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}, \qquad i \in \overline{J}.$$

Note that $k_{J+1} > 0$ holds if $k_i < b_i$ for some $i \in \overline{J}$.

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

Proposition 3.3.1. *There exists a strictly positive measure $\widetilde{\theta} = (\widetilde{\theta}(\mathbf{k}) : \mathbf{k} \in K)$, which will be provided below, such that the measure $\mathbf{x} := (x\left(\mathbf{n},\mathbf{k}\right) : (\mathbf{n},\mathbf{k}) \in E)$ with*

$$x\left(\mathbf{n},\mathbf{k}\right) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}\left(\mathbf{k}\right), \tag{3.3.1}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J},$$

solves the global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ and is therefore stationary for Z. Consequently, $\mathbf{x}$ is strictly positive.

Proof. **The global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$** of the stochastic queueing-inventory process Z are given for $(\mathbf{n},\mathbf{k}) \in E$ by

$$\begin{aligned}
&x\left(\mathbf{n},\mathbf{k}\right) \cdot \Big(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{k_i>0\}} + \sum_{i\in\overline{J}} \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{k_i>0\}} + \sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}\Big)\\
&= \sum_{i\in\overline{J}} x\left(\mathbf{n}-\mathbf{e}_i,\mathbf{k}\right) \cdot \lambda_i \cdot 1_{\{n_i>0\}} \cdot 1_{\{k_i>0\}}\\
&\quad + \sum_{i\in\overline{J}} x\left(\mathbf{n}+\mathbf{e}_i,\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1}\right) \cdot \mu_i(n_i+1) \cdot 1_{\{k_i<b_i\}}\\
&\quad + \sum_{i\in\overline{J}} x\left(\mathbf{n},\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}\right) \cdot \nu \cdot p_i(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}) \cdot 1_{\{k_i>0\}}.
\end{aligned}$$

It has to be shown that the stationary measure from Proposition 3.3.1 satisfies these global balance equations. $\left(\prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j)\right)$ can be separated analogously as shown in the proof of Proposition 2.3.1 on page 15. Consequently, it holds

$$\begin{aligned}
&\widetilde{\theta}\left(\mathbf{k}\right) \Big(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{k_i>0\}} + \sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}\Big)\\
&= \sum_{i\in\overline{J}} \widetilde{\theta}\left(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1}\right) \cdot \lambda_i \cdot 1_{\{k_i<b_i\}}\\
&\quad + \sum_{i\in\overline{J}} \widetilde{\theta}\left(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}\right) \cdot \nu \cdot p_i(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}) \cdot 1_{\{k_i>0\}}.
\end{aligned} \tag{3.3.2}$$

An inspection of the system (3.3.2) reveals that it is a "generator equation", i.e. the global balance equation $\widetilde{\theta} \cdot \mathbf{Q}_{red} = 0$ for a suitably defined ergodic Markov process on state space K with "reduced generator" $\mathbf{Q}_{red} = \left(q_{red}(\mathbf{k}; \widetilde{\mathbf{k}}) : \mathbf{k}, \widetilde{\mathbf{k}} \in K\right)$ with the following transition rates for $\mathbf{k} \in K$:

$$q_{red}\left(\mathbf{k}; \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}\right) = \lambda_i \cdot 1_{\{k_i > 0\}}, \qquad i \in \overline{J},$$
$$q_{red}\left(\mathbf{k}; \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}\right) = \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}}, \qquad i \in \overline{J}.$$

The Markov process generated by $\mathbf{Q}_{red}$ is irreducible on K and therefore (3.3.2) has a solution which is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Remark 3.3.2. In Section 3.3.1.1, the marginal measure $\widetilde{\theta}$ is derived in explicit form for the special case with base stock levels $b_j = 1$, $j \in \overline{J}$.

For systems with two locations and base stock levels greater than one ($b_j > 1$, $j \in \overline{J}$), the marginal measure $\widetilde{\theta}$ can be obtained by a recursive method which is described by the algorithm given in Appendix C.1.

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 3.3.3. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := \left(\pi\left(\mathbf{n}, \mathbf{k}\right) : (\mathbf{n}, \mathbf{k}) \in E\right), \quad \pi\left(\mathbf{n}, \mathbf{k}\right) := \lim_{t \to \infty} P\left(Z(t) = (\mathbf{n}, \mathbf{k})\right)$$

and the appropriate marginal distributions

$$\xi := \left(\xi\left(\mathbf{n}\right) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi\left(\mathbf{n}\right) := \lim_{t \to \infty} P\left((X_1(t), \ldots, X_J(t)) = \mathbf{n}\right),$$
$$\theta := \left(\theta\left(\mathbf{k}\right) : \mathbf{k} \in K\right), \quad \theta\left(\mathbf{k}\right) := \lim_{t \to \infty} P\left((Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) = \mathbf{k}\right).$$

Theorem 3.3.4. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi\left(\mathbf{n}, \mathbf{k}\right) = \xi(\mathbf{n}) \cdot \theta\left(\mathbf{k}\right), \tag{3.3.3}$$

with

$$\xi(\mathbf{n}) = \prod_{j \in \overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{3.3.4}$$

and normalisation constants

$$C_j = \sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}$$

and θ is the probabilistic solution of (3.3.2).

Proof. Z is ergodic if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ from Proposition 3.3.1 can be normalised (i.e. $\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k}) < \infty$). Because of Proposition 3.3.1 it holds

$$\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k}) = \sum_{\mathbf{n}\in\mathbb{N}_0} \widetilde{\xi}(\mathbf{n}) \cdot \sum_{\mathbf{k}\in K} \widetilde{\theta}(\mathbf{k}) = \left(\prod_{j\in\overline{J}} \sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} \right) \cdot \sum_{\mathbf{k}\in K} \widetilde{\theta}(\mathbf{k}) .$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 3.3.1 can be normalised if and only if $\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty$ for all $j \in \overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n},\mathbf{k}) = \frac{x(\mathbf{n},\mathbf{k})}{\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k})},$$

where $x(\mathbf{n},\mathbf{k})$ is given in Proposition 3.3.1. □

Remark 3.3.5. The expression (3.3.3) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

Representation (3.3.4) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j) = \prod_{j\in\overline{J}} C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. In our production network, however, servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (3.3.3) has been unexpected to us.

Our production-inventory system can be considered as a "Jackson network in a random environment" in [KDO16, Section 4]. We can interpret the inventory-replenishment subsystem, which contributes via θ to Theorem 3.3.4, as a "random environment" for the production network of nodes $\overline{J}$, which is a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1] we conclude from the hindsight that decoupling of the queueing process $(X_1, \ldots, X_J)$ and the process $(Y_1, \ldots, Y_J, W_{J+1})$, i.e. the formula (3.3.3), is a consequence of that Theorem 4.1.

Our direct proof of Proposition 3.3.1 is much shorter than embedding the present model into the general framework of [KDO16].

The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

3.3.1.1. Calculation of $\widetilde{\theta}$ for the special case $b_j = 1,\ j \in \overline{J}$

In this section, we will solve the global balance equation $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$ for the special case with base stock levels $b_j = 1,\ j \in \overline{J}$.

We recall the notation for the inventory-replenishment subsystem

$$\widetilde{\theta}(\overbrace{k_1, k_2, \dots, k_J}^{\substack{\text{inventories} \\ \text{at locations}}},\ \overbrace{k_{J+1}}^{\text{supplier}}).$$

Proposition 3.3.6. *The strictly positive measure* $\widetilde{\theta} = \left(\widetilde{\theta}(\mathbf{k}) : \mathbf{k} \in K\right)$ *of the inventory-replenishment subsystem with base stock levels* $b_1 = b_2 = \cdots = b_J = 1$ *is given by*

$$\widetilde{\theta}(\mathbf{k}) = \widetilde{\theta}(k_1, k_2, \dots, k_J, k_{J+1}) = \left(\prod_{j \in A_\mathbf{k}} \frac{\nu}{\lambda_j}\right) \cdot \left(\prod_{\ell=|A_\mathbf{k}^c|+1}^{J} \frac{1}{\ell}\right), \tag{3.3.5}$$

where $A_\mathbf{k} := \left\{i \in \overline{J} | k_i = 1\right\}$ *and* $A_\mathbf{k}^c := \left\{i \in \overline{J} | k_i = 0\right\}$.

Proof. It has to be shown that the stationary measure (3.3.5) satisfies the global balance equations $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$, which are given for $\mathbf{k} \in K$ by

$$\begin{aligned}
&\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}}\Big) \\
&= \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i \cdot 1_{\{k_i < b_i\}} \\
&\quad + \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot 1_{\{k_i > 0\}}
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad &\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i = 1\}} + \sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i = 0\}}\Big) \\
&= \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i \cdot 1_{\{k_i = 0\}} \\
&\quad + \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot 1_{\{k_i = 1\}}
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad &\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i \in A_\mathbf{k}} \lambda_i + \sum_{i \in A_\mathbf{k}^c} \nu \cdot p_i(\mathbf{k})\Big) \\
&= \sum_{i \in A_\mathbf{k}^c} \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i \\
&\quad + \sum_{i \in A_\mathbf{k}} \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}).
\end{aligned}$$

It should be noted that for $i \in A_\mathbf{k}$ holds $A_\mathbf{k} \cup \{i\} = A_{\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1}}$ and for $i \in A_\mathbf{k}$ holds $A_\mathbf{k} \setminus \{i\} = A_{\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}}$.

Substitution of (3.3.5) into the global balance equations directly leads for $\mathbf{k}$ with $k_1 = k_2 = \cdots = k_J = 0$, i.e. $|A_{\mathbf{k}}| = 0$ and $|A_{\mathbf{k}}^c| = J$, to

$$\widetilde{\theta}\,(0,0,\ldots,0,0,J) \cdot \underbrace{\Big(\sum_{i \in A_{\mathbf{k}}^c} \nu \cdot \underbrace{p_i\,(0,0,\ldots,0,0,J)}_{=\frac{1}{|A_{\mathbf{k}}^c|}=\frac{1}{J}} \Big)}_{=\nu} = \sum_{i \in A_{\mathbf{k}}^c} \widetilde{\theta}\,(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i$$

$$\Leftrightarrow \quad \widetilde{\theta}\,(0,0,\ldots,0,0,J) \cdot \nu = \sum_{i \in A_{\mathbf{k}}^c} \widetilde{\theta}\,(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i$$

$$\Leftrightarrow \quad \underbrace{\left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right)}_{=1} \cdot \underbrace{\left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right)}_{=1} \cdot \nu = \sum_{i \in A_{\mathbf{k}}^c} \left(\prod_{j \in A_{\mathbf{k}} \cup \{i\}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c \setminus \{i\}|+1}^{J} \frac{1}{\ell} \right) \cdot \lambda_i$$

$$\Leftrightarrow \quad \nu = \sum_{i \in A_{\mathbf{k}}^c} \underbrace{\left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right)}_{=1} \cdot \underbrace{\left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right)}_{=1} \cdot \frac{\nu}{\lambda_i} \cdot \frac{1}{|A_{\mathbf{k}}^c|} \cdot \lambda_i$$

$$= \sum_{i \in A_{\mathbf{k}}^c} \frac{1}{|A_{\mathbf{k}}^c|} \cdot \nu = |A_{\mathbf{k}}^c| \cdot \frac{1}{|A_{\mathbf{k}}^c|} \cdot \nu = \nu,$$

for $\mathbf{k}$ with $k_1 = k_2 = \cdots = k_J = 1$, i.e. $|A_{\mathbf{k}}| = J$ and $|A_{\mathbf{k}}^c| = 0$, to

$$\widetilde{\theta}\,(1,1,\ldots,1,1,0) \cdot \sum_{i \in A_{\mathbf{k}}} \lambda_i = \sum_{i \in A_{\mathbf{k}}} \widetilde{\theta}\,(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot \underbrace{p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1})}_{=1}$$

$$\Leftrightarrow \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \sum_{i \in A_{\mathbf{k}}} \lambda_i = \sum_{i \in A_{\mathbf{k}}} \left(\prod_{j \in A_{\mathbf{k}} \setminus \{i\}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c \cup \{i\}|+1}^{J} \frac{1}{\ell} \right) \cdot \nu$$

$$\Leftrightarrow \qquad = \sum_{i \in A_{\mathbf{k}}} \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \frac{\lambda_i}{\nu} \cdot \frac{1}{1} \cdot \nu$$

$$= \sum_{i \in A_{\mathbf{k}}} \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \lambda_i,$$

for $0 < |A_{\mathbf{k}}| < J$ and $0 < |A_{\mathbf{k}}^c| < J$ to

$$\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i \in A_{\mathbf{k}}} \lambda_i + \underbrace{\sum_{i \in A_{\mathbf{k}}^c} \nu \cdot \underbrace{p_i(\mathbf{k})}_{=\frac{1}{|A_{\mathbf{k}}^c|}}}_{=\nu} \Big)$$

$$= \sum_{i \in A_{\mathbf{k}}^c} \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i + \sum_{i \in A_{\mathbf{k}}} \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot \underbrace{p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1})}_{=\frac{1}{|A_{\mathbf{k}}^c|+1}}$$

$$\Leftrightarrow \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \Big(\sum_{i \in A_{\mathbf{k}}} \lambda_i + \nu \Big)$$

$$= \sum_{i \in A_{\mathbf{k}}^c} \left(\prod_{j \in A_{\mathbf{k}} \cup \{i\}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c \setminus \{i\}|+1}^{J} \frac{1}{\ell} \right) \cdot \lambda_i + \sum_{i \in A_{\mathbf{k}}} \left(\prod_{j \in A_{\mathbf{k}} \setminus \{i\}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c \cup \{i\}|+1}^{J} \frac{1}{\ell} \right) \cdot \nu \cdot \frac{1}{|A_{\mathbf{k}}^c| + 1}$$

$$= \sum_{i \in A_{\mathbf{k}}^c} \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \frac{\nu}{\lambda_i} \cdot \frac{1}{|A_{\mathbf{k}}^c|} \cdot \lambda_i + \sum_{i \in A_{\mathbf{k}}} \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \frac{\lambda_i}{\nu} \cdot (|A_{\mathbf{k}}^c| + 1) \cdot \nu \cdot \frac{1}{|A_{\mathbf{k}}^c| + 1}$$

$$= \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \nu + \left(\prod_{j \in A_{\mathbf{k}}} \frac{\nu}{\lambda_j} \right) \cdot \left(\prod_{\ell=|A_{\mathbf{k}}^c|+1}^{J} \frac{1}{\ell} \right) \cdot \sum_{i \in A_{\mathbf{k}}} \lambda_i.$$

□

Remark 3.3.7. We make a distinction between homogeneous and heterogeneous locations.

We mean by *homogeneous* locations that the inventories have identical base stock levels $b_1 = b_2 = \cdots = b_J$ and identical arrival rates $\lambda_1 = \lambda_2 = \cdots = \lambda_J > 0$. Service rates $\mu_j(\cdot) > 0$, $j \in \overline{J}$, obey no such restrictions.

We mean by *heterogeneous* locations that there may be different arrival rates $\lambda_j > 0$, $j \in \overline{J}$ (and any service rate $\mu_j(\cdot) > 0$, $j \in \overline{J}$) and for the base stock levels hold $b_1 \geq b_2 \geq \cdots \geq b_J$.

As a consequence of the preceding Proposition 3.3.6 the following *symmetry property* for homogeneous locations with base stock levels $b_1 = b_2 = \cdots = b_J = 1$ is valid.

For all permutations σ of $\{1, \ldots, J\}$ holds

$$\widetilde{\theta}(\overbrace{k_1, k_2, \ldots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{k_{J+1}}^{\text{supplier}}) = \widetilde{\theta}(\overbrace{k_{\sigma(1)}, k_{\sigma(2)}, \ldots, k_{\sigma(J)}}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{k_{\sigma J+1}}^{\text{supplier}}).$$

For $b_1 = b_2 = \cdots = b_J > 1$ the global balance equations (3.3.2) reveal directly that this symmetry property holds in this case as well.

3.3.1.2. Structural properties of the stationary inventory-replenishment subsystem

In this section, we assume that the queueing-inventory process Z is ergodic. We make again a distinction between homogeneous and heterogeneous locations.

In this section, we will use an abbreviated notation because $k_{J+1} = \sum_{j=1}^{J}(b_j - k_j)$ and the base stock levels b_j, $j \in \overline{J}$, are fixed parameters:

$$\theta(\overbrace{k_1, k_2, \ldots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}}) := \theta(\overbrace{k_1, k_2, \ldots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{k_{J+1}}^{\text{supplier}}).$$

Homogeneous locations

In this paragraph, we assume that the locations are homogeneous, i.e. the inventories have identical base stock levels $b_1 = b_2 = \cdots = b_J > 1$ and identical arrival rates $\lambda_1 = \lambda_2 = \cdots = \lambda_J > 0$. We derived in Proposition 3.3.6 an explicit solution for the special case with base stock levels $b_1 = b_2 = \cdots = b_J = 1$.

Recall that the following *symmetry property* for homogeneous locations holds: For all permutations σ of $\{1, \ldots, J\}$ holds

$$\theta(\overbrace{k_1, k_2, \ldots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}}) = \theta(\overbrace{k_{\sigma(1)}, k_{\sigma(2)}, \ldots, k_{\sigma(J)}}^{\substack{\text{inventories}\\ \text{at locations}}}).$$

Let $(Y_1, Y_2, \dots Y_J, W_{J+1})$ be a random variable which is distributed according to the marginal steady state probability for the inventory-replenishment subsystem.[2]

Proposition 3.3.8. *For the inventory process holds*

$$P(Y_1 = \ell) = P(Y_j = \ell), \quad j \in \overline{J} \setminus \{1\}, \ \ell \in \{0, \dots, b_1\},$$

and

$$\begin{aligned} & P(Y_1 = \ell) \cdot \lambda_1 \\ = {} & P\left(Y_j = \ell - 1 \text{ for } j \in \overline{J}\right) \cdot \frac{\nu}{J} \\ & + \sum_{i=1}^{J-1} P\left(Y_j = \ell - 1 \text{ for } j = 1, \dots, i \text{ and } \ell - 1 < Y_k \leq b_k \text{ for } k = i+1, \dots, J\right) \\ & \quad \cdot \binom{J-1}{i-1} \cdot \frac{\nu}{i}, \qquad \ell \in \{1, \dots, b_1\}. \end{aligned} \tag{3.3.6}$$

Proof. It holds

$$P(Y_1 = \ell) = P(Y_j = \ell), \quad j \in \overline{J} \setminus \{1\}, \ \ell \in \{0, \dots, b_1\}$$

because of the symmetry property for homogeneous locations

$$\theta(k_1, k_2, \dots, k_J) = \theta(k_2, k_1, \dots, k_J) = \cdots = \theta(k_J, k_{J-1}, \dots, k_1)$$

and because

$$P(Y_1 = \ell) = \sum_{k_2=0}^{b_2} \cdots \sum_{k_J=0}^{b_J} \theta(\ell, k_2, \dots, k_J).$$

The equation (3.3.6) can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a) on page 259.

For $\ell \in \{1, \dots, b_1\}$, it can be proven by a cut, which divides E into complementary sets according to the size of the inventory at location 1 that is less than or equal to $\ell - 1$ or greater than $\ell - 1$, i.e. into the sets

$$\Big\{(k_1, k_2, \dots, k_J) : k_1 \in \{0, 1, \dots, \ell{-}1\}, k_j \in \{0, \dots, b_j\}, \ j \in \{2, \dots, J\}\Big\},$$

$$\Big\{(\widetilde{k}_1, \widetilde{k}_2, \dots, \widetilde{k}_J) : \widetilde{k}_1 \in \{\ell, \dots, b_1\}, \ \widetilde{k}_j \in \{0, \dots, b_j\}, \ j \in \{2, \dots, J\}\Big\}, \qquad \ell \in \{1, \dots, b_1\}.$$

[2] It should be noted that $\theta(k_1, k_2, \dots, k_J, k_{J+1}) = P(Y_1 = k_1, Y_2 = k_2, \dots, Y_J = k_J, W_{J+1} = k_{J+1}) = P(Y_1 = k_1, Y_2 = k_2, \dots, Y_J = k_J)$ because the base stock levels b_j, $j \in \overline{J}$, are fixed parameters and $k_{J+1} = \sum_{j=1}^{J} (b_j - k_j)$.

Then, it follows for $\ell \in \{1, \dots, b_1\}$

$$\sum_{k_1=0}^{\ell-1} \sum_{k_2=0}^{b_2} \cdots \sum_{k_J=0}^{b_J} \sum_{\widetilde{k}_1=\ell}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \cdots \sum_{\widetilde{k}_J=0}^{b_J} \theta(k_1, k_2, \dots, k_J) \cdot q_{red}\Big((k_1, k_2, \dots, k_J); (\widetilde{k}_1, \widetilde{k}_2, \dots, \widetilde{k}_J)\Big)$$

$$= \sum_{\widetilde{k}_1=\ell}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \cdots \sum_{\widetilde{k}_J=0}^{b_J} \sum_{k_1=0}^{\ell-1} \sum_{k_2=0}^{b_2} \cdots \sum_{k_J=0}^{b_J} \theta(\widetilde{k}_1, \widetilde{k}_2, \dots, \widetilde{k}_J) \cdot q_{red}\Big((\widetilde{k}_1, \widetilde{k}_2, \dots, \widetilde{k}_J); (k_1, k_2, \dots, k_J)\Big)$$

$$\Leftrightarrow \quad \theta(\ell-1, \ell-1, \dots, \ell-1) \cdot \frac{\nu}{J} \tag{3.3.7}$$

$$+\sum_{i=1}^{J-1} \sum_{k_{i+1}=\ell}^{b_{i+1}} \cdots \sum_{k_J=\ell}^{b_J} \theta(\ell-1, \ell-1, \dots, \ell-1, k_{i+1}, \dots, k_J) \cdot \binom{J-1}{i-1} \cdot \frac{1}{i} \cdot \nu \tag{3.3.8}$$

$$= \sum_{\widetilde{k}_2=0}^{b_2} \cdots \sum_{\widetilde{k}_J=0}^{b_J} \theta(\ell, \widetilde{k}_2, \dots, \widetilde{k}_J) \cdot \lambda_1.$$

The only possible transitions from the set, where the size of the inventory at location 1 is less than or equal to $\ell - 1$, to the set, where the inventory at location 1 is greater than $\ell - 1$, are transitions according to a replenishment. In particular, transitions from

$$\Big\{(\ell-1, k_2, \dots, k_J) : k_j \in \{\ell-1, \dots, b_j\},\ j \in \{2, \dots, J\}\Big\},$$

to

$$\Big\{(\ell, \widetilde{k}_2, \dots, \widetilde{k}_J) : \widetilde{k}_j \in \{\ell-1, \dots, b_j\},\ j \in \{2, \dots, J\}\Big\}, \qquad \ell \in \{1, \dots, b_1\}.$$

A replenishment at location 1 is only possible if $\{1\} \subseteq \arg\max_{j \in \overline{J}}(b_j - k_j)$. This means that there is no other location with higher difference between the on-hand inventory and the capacity of the inventory (= base stock level). Consequently, all possible states where the other locations have $\ell - 1$ items or more items in the inventory have to be considered. In (3.3.7) all locations have exactly $\ell - 1$ items in the inventory and in (3.3.8) i states how many locations have exactly $\ell - 1$ items in the inventory. This results in the factor $1/i$, which is the probability that the finished item is sent to location i. The symmetry property leads to the factor $\binom{J-1}{i-1}$.

Hence, we have shown for $\ell \in \{1, \dots, b_1\}$

$$P(Y_1 = \ell) \cdot \lambda_1$$

$$= P\left(Y_j = \ell - 1 \text{ for } j \in \overline{J}\right) \cdot \frac{\nu}{J}$$

$$+\sum_{i=1}^{J-1} P\left(Y_j = \ell - 1 \text{ for } j = 1, \dots, i \text{ and } \ell - 1 < Y_k \leq b_k \text{ for } k = i+1, \dots, J\right) \cdot \binom{J-1}{i-1} \cdot \frac{\nu}{i}.$$

□

Heterogeneous locations
In this paragraph, we assume that there are two heterogeneous locations with base stock levels $b_1 \geq b_2$, where $b_1 > 1$ and $b_2 \geq 1$ and arrival rates $\lambda_1, \lambda_2 > 0$.

Let (Y_1, Y_2, W_3) be a random variable which is distributed according to the marginal steady state probability for the inventory-replenishment subsystem.[3]

Remark 3.3.9. From equation (3.3.9) in the following proposition follows

$$P(Y_1 = \ell_1) = P(Y_1 = 0) \cdot \left(\frac{\nu}{\lambda_1}\right)^{\ell_1}, \qquad \ell_1 \in \{1, \dots, b_1 - b_2\}.$$

Proposition 3.3.10. *For the inventory process holds for* $\ell_1 = 1, \dots, b_1 - b_2$

$$P(Y_1 = \ell_1) \cdot \lambda_1 = P(Y_1 = \ell_1 - 1) \cdot \nu, \tag{3.3.9}$$

for $\ell_1 = b_1 - b_2 + 1, \dots, b_1 - 1$

$$\begin{aligned} P(Y_1 = \ell_1) \cdot \lambda_1 = P(Y_1 = \ell_1 - 1, Y_2 = \ell_1 - 1 - b_1 + b_2) \cdot \frac{1}{2}\nu \\ + P(Y_1 = \ell_1 - 1, Y_2 > \ell_1 - 1 - b_1 + b_2) \cdot \nu, \end{aligned} \tag{3.3.10}$$

for $\ell_1 = b_1$

$$\begin{aligned} P(Y_1 = b_1) \cdot \lambda_1 = P(Y_1 = b_1 - 1, Y_2 = b_2 - 1) \cdot \frac{1}{2}\nu \\ + P(Y_1 = b_1 - 1, Y_2 = b_2) \cdot \nu, \end{aligned} \tag{3.3.11}$$

for $\ell_2 = 1, \dots, b_2$

$$\begin{aligned} P(Y_2 = \ell_2) \cdot \lambda_2 = P(Y_1 = b_1 - b_2 + \ell_2 - 1, Y_2 = \ell_2 - 1) \cdot \frac{1}{2}\nu \\ + P(Y_1 > b_1 - b_2 + \ell_2 - 1, Y_2 = \ell_2 - 1) \cdot \nu. \end{aligned} \tag{3.3.12}$$

Proof. The equations can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a) on page 259.

For $\ell_1 \in \{1, \dots, b_1 - b_2\}$, equation (3.3.9) can be proven by a cut, which divides E into complementary sets according to the size of the inventory at location 1 that is less than or equal to $\ell_1 - 1$ or greater than $\ell_1 - 1$, i.e. into the sets

$$\Big\{(k_1, k_2) : k_1 \in \{0, 1, \dots, \ell_1 - 1\},\ k_2 \in \{0, \dots, b_2\}\Big\},$$

$$\Big\{(\widetilde{k}_1, \widetilde{k}_2) : \widetilde{k}_1 \in \{\ell_1, \dots, b_1\},\ \widetilde{k}_2 \in \{0, \dots, b_2\}\Big\}, \quad \ell_1 \in \{1, \dots, b_1 - b_2\}.$$

[3]It should be noted that $\theta(k_1, k_2, k_3) = P(Y_1 = k_1, Y_2 = k_2, W_3 = k_3) = P(Y_1 = k_1, Y_2 = k_2)$, because the base stock levels b_1 and b_2 are fixed parameters and $k_3 = (b_1 + b_2) - (k_1 + k_2)$.

Then, it follows for $\ell_1 \in \{1, \ldots, b_1 - b_2\}$

$$\begin{aligned}
&\sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \theta(k_1, k_2) \cdot q_{red}\left((k_1, k_2); (\widetilde{k}_1, \widetilde{k}_2)\right) \\
= &\sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \theta(\widetilde{k}_1, \widetilde{k}_2) \cdot q_{red}\left((\widetilde{k}_1, \widetilde{k}_2); (k_1, k_2)\right) \\
\Leftrightarrow \quad &\underbrace{\sum_{k_2=0}^{b_2} \theta(\ell_1 - 1, k_2)}_{=P(Y_1=\ell_1-1)} \cdot \nu = \underbrace{\sum_{\widetilde{k}_2=0}^{b_2} \theta(\ell_1, \widetilde{k}_2)}_{=P(Y_1=\ell_1)} \cdot \lambda_1.
\end{aligned}$$

Hence, we have shown for $\ell_1 \in \{1, \ldots, b_1 - b_2\}$

$$P(Y_1 = \ell_1 - 1) \cdot \nu = P(Y_1 = \ell_1) \cdot \lambda_1.$$

For $\ell_1 \in \{b_1 - b_2 + 1, \ldots, b_1 - 1\}$, equation (3.3.10) can be proven by a cut, which divides E into complementary sets according to the size of the inventory at location 1 that is less than or equal to $\ell_1 - 1$ or greater than $\ell_1 - 1$, i.e. into the sets

$$\begin{aligned}
&\{(k_1, k_2) : k_1 \in \{0, 1, \ldots, \ell_1 - 1\},\ k_2 \in \{0, \ldots, b_2\}\}, \\
&\left\{(\widetilde{k}_1, \widetilde{k}_2) : \widetilde{k}_1 \in \{\ell_1, \ldots, b_1\},\ \widetilde{k}_2 \in \{0, \ldots, b_2\}\right\}, \quad \ell_1 \in \{b_1 - b_2 + 1, \ldots, b_1 - 1\}.
\end{aligned}$$

Then, it follows for $\ell_1 \in \{b_1 - b_2 + 1, \ldots, b_1 - 1\}$

$$\begin{aligned}
&\sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \theta(k_1, k_2) \cdot q_{red}\left((k_1, k_2); (\widetilde{k}_1, \widetilde{k}_2)\right) \\
= &\sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \theta(\widetilde{k}_1, \widetilde{k}_2) \cdot q_{red}\left((\widetilde{k}_1, \widetilde{k}_2); (k_1, k_2)\right) \\
\Leftrightarrow \quad &\underbrace{\sum_{k_2=b_2-(b_1-\ell_1)}^{b_2} \theta(\ell_1 - 1, k_2) \cdot \nu}_{=P(Y_1=\ell_1-1, Y_2>b_2-(b_1-(\ell_1-1)))} + \underbrace{\theta(\ell_1 - 1, b_2 - (b_1 - (\ell_1 - 1))) \cdot \frac{1}{2}\nu}_{=P(Y_1=\ell_1-1, Y_2=b_2-(b_1-(\ell_1-1)))} = \underbrace{\sum_{\widetilde{k}_2=0}^{b_2} \theta(\ell_1, \widetilde{k}_2)}_{=P(Y_1=\ell_1)} \cdot \lambda_1.
\end{aligned}$$

Hence, we have shown for $\ell_1 \in \{b_1 - b_2 + 1, \ldots, b_1 - 1\}$

$$\begin{aligned}
P(Y_1 = \ell_1) \cdot \lambda_1 &= P(Y_1 = \ell_1 - 1, Y_2 = b_2 - (b_1 - (\ell_1 - 1))) \cdot \frac{1}{2}\nu \\
&\quad + P(Y_1 = \ell_1 - 1, Y_2 > b_2 - (b_1 - (\ell_1 - 1))) \cdot \nu \\
&= P(Y_1 = \ell_1 - 1, Y_2 = \ell_1 - 1 - b_1 + b_2) \cdot \frac{1}{2}\nu \\
&\quad + P(Y_1 = \ell_1 - 1, Y_2 > \ell_1 - 1 - b_1 + b_2) \cdot \nu.
\end{aligned}$$

For $\ell_1 = b_1$, equation (3.3.11) can be proven by a cut, which divides E into complementary sets according to the size of the inventory at location 1 that is less than or equal to $b_1 - 1$ or greater than $b_1 - 1$, i.e. into the sets

$$\Big\{(k_1, k_2) : k_1 \in \{0, 1, \ldots, b_1 - 1\},\ k_2 \in \{0, \ldots, b_2\}\Big\},$$
$$\Big\{(\widetilde{k}_1, \widetilde{k}_2) : \widetilde{k}_1 = b_1,\ \widetilde{k}_2 \in \{0, \ldots, b_2\}\Big\}.$$

Then, it follows for

$$\begin{aligned}
&\sum_{k_1=0}^{b_1-1} \sum_{k_2=0}^{b_2} \sum_{\widetilde{k}_1=b_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \theta(k_1, k_2) \cdot q_{red}\Big((k_1, k_2); (\widetilde{k}_1, \widetilde{k}_2)\Big) \\
&= \sum_{\widetilde{k}_1=b_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \sum_{k_1=0}^{b_1-1} \sum_{k_2=0}^{b_2} \theta(\widetilde{k}_1, \widetilde{k}_2) \cdot q_{red}\Big((\widetilde{k}_1, \widetilde{k}_2); (k_1, k_2)\Big) \\
\Leftrightarrow\quad & \underbrace{\theta(b_1 - 1, b_2)}_{=P(Y_1=b_1-1, Y_2=b_2)} \cdot \nu + \underbrace{\theta(b_1 - 1, b_2 - 1)}_{=P(Y_1=b_1-1, Y_2=b_2-1)} \cdot \frac{1}{2}\nu = \underbrace{\sum_{\widetilde{k}_2=0}^{b_2} \theta(b_1, \widetilde{k}_2)}_{=P(Y_1=b_1)} \cdot \lambda_1.
\end{aligned}$$

Hence, we have shown

$$\begin{aligned}
P(Y_1 = b_1) \cdot \lambda_1 = {} & P(Y_1 = b_1 - 1, Y_2 = b_2 - 1) \cdot \frac{1}{2}\nu \\
& + P(Y_1 = b_1 - 1, Y_2 = b_2) \cdot \nu.
\end{aligned}$$

For $\ell_2 \in \{1, \ldots, b_2\}$, equation (3.3.12) can be proven by a cut, which divides E into complementary sets according to the size of the inventory at location 2 that is less than or equal to $\ell_2 - 1$ or greater than $\ell_2 - 1$, i.e. into the sets

$$\Big\{(k_1, k_2) : k_1 \in \{0, \ldots, b_1\},\ k_2 \in \{0, 1, \ldots, \ell_2 - 1\}\Big\},$$
$$\Big\{(\widetilde{k}_1, \widetilde{k}_2) : \widetilde{k}_1 \in \{0, \ldots, b_1\},\ \widetilde{k}_2 \in \{\ell_2, \ldots, b_2\}\Big\}, \quad \ell_2 \in \{1, \ldots, b_2\}.$$

Then, it follows for $\ell_2 \in \{1, \ldots, b_2\}$

$$\begin{aligned}
&\sum_{k_1=0}^{b_1} \sum_{k_2=0}^{\ell_2-1} \sum_{\widetilde{k}_1=0}^{b_1} \sum_{\widetilde{k}_2=\ell_2}^{b_2} \theta(k_1, k_2) \cdot q_{red}\Big((k_1, k_2); (\widetilde{k}_1, \widetilde{k}_2)\Big) \\
&= \sum_{\widetilde{k}_1=0}^{b_1} \sum_{\widetilde{k}_2=\ell_2}^{b_2} \sum_{k_1=0}^{b_1} \sum_{k_2=0}^{\ell_2-1} \theta(\widetilde{k}_1, \widetilde{k}_2) \cdot q_{red}\Big((\widetilde{k}_1, \widetilde{k}_2); (k_1, k_2)\Big) \\
\Leftrightarrow\quad & \underbrace{\theta(b_1 - (b_2 - (\ell_2 - 1)), \ell_2 - 1) \cdot \frac{1}{2}\nu}_{=P(Y_1=b_1-(b_2-(\ell_2-1)), Y_2=\ell_2-1)} \quad \underbrace{\sum_{k_1=b_1-(b_2-\ell_2)}^{b_1} \theta(k_1, \ell_2 - 1)}_{=P(Y_1>b_1-(b_2-(\ell_2-1)), Y_2=\ell_2-1)} \cdot \nu = \underbrace{\sum_{\widetilde{k}_1=0}^{b_1} \theta(\widetilde{k}_1, \ell_2)}_{=P(Y_2=\ell_2)} \cdot \lambda_2.
\end{aligned}$$

Hence, we have shown for $\ell_2 \in \{1, \dots, b_2\}$

$$\begin{aligned} P(Y_2 = \ell_2) \cdot \lambda_2 &= P(Y_1 = b_1 - (b_2 - (\ell_2 - 1)), Y_2 = \ell_2 - 1) \cdot \frac{1}{2}\nu \\ &\quad + P(Y_1 > b_1 - (b_2 - (\ell_2 - 1)), Y_2 = \ell_2 - 1) \cdot \nu \\ &= P(Y_1 = b_1 - b_2 + \ell_2 - 1, Y_2 = \ell_2 - 1) \cdot \frac{1}{2}\nu \\ &\quad + P(Y_1 > b_1 - b_2 + \ell_2 - 1, Y_2 = \ell_2 - 1) \cdot \nu. \end{aligned}$$

□

3.4. Load balancing policy: Weak priorities

In this section, we study the supply chain with weak priorities for load balancing policy as described in Section 3.2.

We assume that the finished item of raw material is sent with probability (cf. [Tow80, Section 3, Example 1, pp. 326ff.] and [Dad85, pp. 624f.])

$$p_j(\mathbf{k}) = h_j\left(k_j\right) \cdot h\left(\sum_{i \in \overline{J}} k_i\right) > 0$$

to location $j \in \overline{J}$, if the inventory is not full at this location (this means that the on-hand inventory level at location j is lower than the base stock level b_j).

We define

$$h_j : \{0, 1, \ldots, b_j\} \to \{0, 1, \ldots, b_j\}, \qquad h_j(k_j) = b_j - k_j, \qquad j \in \overline{J},$$

and

$$h : \left\{0, 1, \ldots, \sum_{j \in \overline{J}} b_j - 1\right\} \to \mathbb{R}^+, \qquad h(m) = \left(\sum_{j \in \overline{J}} b_j - m\right)^{-1}.$$

Then, it holds

$$h\left(\sum_{i \in \overline{J}} k_i\right) = \left(\sum_{i \in \overline{J}} b_i - \sum_{i \in \overline{J}} k_i\right)^{-1} = \left(k_{J+1}\right)^{-1}.$$

The routing probabilities out of the central supplier must sum to one if there is at least one order at the central supplier. The finished item is sent with greater probability to the location with higher difference between the on-hand inventory and the capacity of the inventory.

3.4.1. Limiting and stationary distribution

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z; \tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n}, \mathbf{k}) \in E$:

$$\begin{aligned}
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n} + \mathbf{e}_i, \mathbf{k})\right) &= \lambda_i \cdot 1_{\{k_i > 0\}}, && i \in \overline{J}, \\
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n} - \mathbf{e}_i, \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1})\right) &= \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}}, && i \in \overline{J}, \\
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n}, \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1})\right) &= \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}}, && i \in \overline{J},
\end{aligned}$$

with $p_i(\mathbf{k}) = h_i(k_i) \cdot h\left(\sum_{j \in \overline{J}} k_j\right)$. Note that $k_{J+1} > 0$ holds if $k_i < b_i$ for some $i \in \overline{J}$.

Furthermore, $q(z; \tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z; z\right) = -\sum_{\substack{\tilde{z} \in E, \\ z \neq \tilde{z}}} q\left(z; \tilde{z}\right) \qquad \forall z \in E.$$

Proposition 3.4.1. *The strictly positive measure* $\mathbf{x} := (x(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E)$ *with*

$$x(\mathbf{n},\mathbf{k}) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}(\mathbf{k}), \tag{3.4.1}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{3.4.2}$$

$$\widetilde{\theta}(\mathbf{k}) = \left(\prod_{j\in\overline{J}}\left(\prod_{\ell=0}^{k_j-1} h_j(\ell)\right)\right) \cdot \left(\prod_{\ell=0}^{\sum_{j\in\overline{J}} k_j-1} h(\ell)\right) \cdot \left(\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{k_j}\right) \cdot \left(\frac{1}{\nu}\right)^{k_{J+1}}, \quad \mathbf{k} \in K, \tag{3.4.3}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z*.*

Proof. **The global balance equations** $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are given for $(\mathbf{n},\mathbf{k}) \in E$ by

$$\begin{aligned}
&x(\mathbf{n},\mathbf{k}) \left(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{k_i>0\}} + \sum_{i\in\overline{J}} \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{k_i>0\}} + \sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}\right)\\
&= \sum_{i\in\overline{J}} x(\mathbf{n}-\mathbf{e}_i,\mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i>0\}} \cdot 1_{\{k_i>0\}}\\
&\quad + \sum_{i\in\overline{J}} x(\mathbf{n}+\mathbf{e}_i,\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1}) \cdot \mu_i(n_i+1) \cdot 1_{\{k_i<b_i\}}\\
&\quad + \sum_{i\in\overline{J}} x(\mathbf{n},\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}) \cdot 1_{\{k_i>0\}}.
\end{aligned}$$

It has to be shown that the stationary measure (3.4.1) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

$\left(\prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j)\right)$ can be separated analogously as shown in the proof of Proposition 2.3.1 on page 15. Consequently, it holds

$$\begin{aligned}
&\widetilde{\theta}(\mathbf{k}) \left(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{k_i>0\}} + \sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}\right)\\
&= \sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1}) \cdot \lambda_i \cdot 1_{\{k_i<b_i\}}\\
&\quad + \sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1}) \cdot 1_{\{k_i>0\}}.
\end{aligned} \tag{3.4.4}$$

Substitution of (3.4.3) we obtain

$$\widetilde{\theta}(\mathbf{k}) \cdot \left(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}} \right)$$

$$= \sum_{i \in \overline{J}} \underbrace{\left(\prod_{j \in \overline{J}} \left(\prod_{\ell=0}^{k_j - 1} h_j(\ell) \right) \right) \cdot \left(\prod_{\ell=0}^{\sum_{j \in \overline{J}} k_j - 1} h(\ell) \right) \cdot \left(\prod_{j \in \overline{J}} \left(\frac{1}{\lambda_j} \right)^{k_j} \right) \cdot \left(\frac{1}{\nu} \right)^{k_{J+1}}}_{=\widetilde{\theta}(\mathbf{k})}$$

$$\cdot \underbrace{h_i(k_i) \cdot h\left(\sum_{j \in \overline{J}} k_j \right)}_{=p_i(\mathbf{k})} \cdot \frac{1}{\lambda_i} \cdot \nu \cdot \lambda_i \cdot 1_{\{k_i < b_i\}}$$

$$+ \sum_{i \in \overline{J}} \underbrace{\left(\prod_{j \in \overline{J}} \left(\prod_{\ell=0}^{k_j - 1} h_j(\ell) \right) \right) \cdot \left(\prod_{\ell=0}^{\sum_{j \in \overline{J}} k_j - 1} h(\ell) \right) \cdot \left(\prod_{j \in \overline{J}} \left(\frac{1}{\lambda_j} \right)^{k_j} \right) \cdot \left(\frac{1}{\nu} \right)^{k_{J+1}}}_{=\widetilde{\theta}(\mathbf{k})}$$

$$\cdot \underbrace{\frac{1}{h_i(k_i - 1)} \cdot \frac{1}{h\left(\sum_{j \in \overline{J}} k_j - 1 \right)}}_{=\frac{1}{p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1})}} \cdot \lambda_i \cdot \frac{1}{\nu} \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot 1_{\{k_i > 0\}}.$$

The right-hand side of the last equation is

$$\sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k}) \cdot \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}} + \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{k_i > 0\}},$$

which is obviously the left-hand side.

Inspection of the system (3.4.4) reveals that it is a "generator equation", i.e. the global balance equation $\widetilde{\theta} \cdot \mathbf{Q}_{red} = 0$ for a suitably defined ergodic Markov process on state space K with "reduced generator" $\mathbf{Q}_{red}$. Because the Markov process generated by $\mathbf{Q}_{red}$ is irreducible the solution of (3.4.4) is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Remark 3.4.2. $\widetilde{\theta}(\mathbf{k}) = \widetilde{\theta}(k_1, \ldots, k_J, k_{J+1})$ is obtained as a strictly positive solution of (3.4.4) which resembles the global balance equations of an artificial non-standard Gordon-Newell network of queues with $J+1$ nodes and $\sum_{j \in \overline{J}} b_j$ customers, exponentially distributed service times with rate λ_j for $k_j \leq b_j$ and "∞" otherwise at node $j \in \{1, \ldots, J\}$ and with rate ν at node $J+1$ and state-dependent routing probabilities. More precisely, it is a starlike system with $r(j, J+1) = 1$, $j \in \overline{J}$, and with state-dependent branching probabilities $r(J+1, j) = p_j(\mathbf{k})$, $j \in \overline{J}$. The $p_j(\mathbf{k}) = h_j(k_j) \cdot h(\sum_{i \in \overline{J}} k_i)$ results from the load balancing policy with weak priorities.

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 3.4.3. For the queueing-inventory system Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E), \quad \pi(\mathbf{n},\mathbf{k}) := \lim_{t\to\infty} P(Z(t) = (\mathbf{n},\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := \left(\xi(\mathbf{n}) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi(\mathbf{n}) := \lim_{t\to\infty} P((X_1(t),\ldots,X_J(t)) = \mathbf{n}),$$

$$\theta := (\theta(\mathbf{k}) : \mathbf{k} \in K), \quad \theta(\mathbf{k}) := \lim_{t\to\infty} P((Y_1(t),\ldots,Y_J(t),W_{J+1}(t)) = \mathbf{k}).$$

Theorem 3.4.4. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n},\mathbf{k}) = \xi(\mathbf{n}) \cdot \theta(\mathbf{k}), \tag{3.4.5}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{3.4.6}$$

$$\theta(\mathbf{k}) = C_\theta^{-1} \cdot \left(\prod_{j\in\overline{J}} \left(\prod_{\ell=0}^{k_j-1} h_j(\ell)\right)\right) \cdot \left(\prod_{\ell=0}^{\sum_{j\in\overline{J}} k_j-1} h(\ell)\right) \cdot \left(\prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{k_j}\right) \cdot \left(\frac{1}{\nu}\right)^{k_{J+1}}, \quad \mathbf{k} \in K, \tag{3.4.7}$$

and normalisation constants

$$C_j = \sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)},$$

$$C_\theta = \sum_{\mathbf{k}\in K} \left(\prod_{j\in\overline{J}} \left(\prod_{\ell=0}^{k_j-1} h_j(\ell)\right)\right) \cdot \left(\prod_{\ell=0}^{\sum_{j\in\overline{J}} k_j-1} h(\ell)\right) \cdot \left(\prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{k_j}\right) \cdot \left(\frac{1}{\nu}\right)^{k_{J+1}}.$$

Remark 3.4.5. Looking at expression (3.4.5), we observe that the corresponding modification of Remark 3.3.5 on page 49 holds here as well.

The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

Remark 3.4.6. Since it holds

$$\prod_{\ell=0}^{\sum_{j\in\overline{J}} k_j-1} h(\ell) = \prod_{\ell=0}^{\sum_{j\in\overline{J}} b_j-k_{J+1}-1} h(\ell) = \prod_{\ell=0}^{\sum_{j\in\overline{J}} b_j-k_{J+1}-1} \frac{1}{\sum_{j\in\overline{J}} b_j-\ell}$$
$$= \frac{1}{\sum_{j\in\overline{J}} b_j}\cdot\frac{1}{\sum_{j\in\overline{J}} b_j-1}\cdots\frac{1}{k_{J+1}+1}$$
$$= \frac{1}{\sum_{j\in\overline{J}} b_j}\cdot\frac{1}{\sum_{j\in\overline{J}} b_j-1}\cdots\frac{1}{k_{J+1}+1}\cdot\underbrace{\frac{k_{J+1}}{k_{J+1}}\cdots\frac{2}{2}\cdot\frac{1}{1}}_{=1}$$
$$= \frac{1}{\left(\sum_{j\in\overline{J}} b_j\right)!}\cdot\prod_{m=0}^{k_{J+1}-1}(m+1), \tag{3.4.8}$$

it follows

$$\theta(\mathbf{k}) \overset{(3.4.7)}{=} C_\theta^{-1}\cdot\left(\prod_{j\in\overline{J}}\left(\prod_{\ell=0}^{k_j-1} h_j(\ell)\right)\right)\cdot\left(\prod_{\ell=0}^{\sum_{j\in\overline{J}} k_j-1} h(\ell)\right)\cdot\left(\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{k_j}\right)\cdot\left(\frac{1}{\nu}\right)^{k_{J+1}}$$
$$\overset{(3.4.8)}{=} C_\theta^{-1}\cdot\left(\prod_{j\in\overline{J}}\left(\prod_{\ell=0}^{k_j-1}\frac{h_j(\ell)}{\lambda_j}\right)\right)\cdot\left(\frac{1}{\left(\sum_{j\in\overline{J}} b_j\right)!}\cdot\prod_{m=0}^{k_{J+1}-1}\frac{m+1}{\nu}\right)$$
$$= \widetilde{C}_\theta^{-1}\cdot\left(\prod_{j\in\overline{J}}\left(\prod_{\ell=0}^{k_j-1}\frac{h_j(\ell)}{\lambda_j}\right)\right)\cdot\left(\prod_{m=0}^{k_{J+1}-1}\frac{m+1}{\nu}\right) \tag{3.4.9}$$

with normalisation constant

$$\widetilde{C}_\theta := \left(\sum_{j\in\overline{J}} b_j\right)!\cdot C_\theta = \sum_{\mathbf{k}\in K}\left(\prod_{j\in\overline{J}}\left(\prod_{\ell=0}^{k_j-1}\frac{h_j(\ell)}{\lambda_j}\right)\right)\cdot\left(\prod_{m=0}^{k_{J+1}-1}\frac{m+1}{\nu}\right).$$

Hence, this explicit formula (3.4.9) for θ shows that in fact there exists a three-term product structure, and that moreover the equilibrium for the integrated model is stratified. In the upper stratum, we have two independent vectors for production and inventory-replenishment, the latter splits into two products, a factor for the subsystem comprising the inventories and a factor for the replenishment subsystem.

In the lower stratum each of the three factors of the upper stratum is decomposed completely in "single-component" factors concerning the production servers, the inventories, and the replenishment servers. It should be noted that the factors for the inventories and the replenishment servers do not indicate internal independence, but they are of product form as the celebrated conditionally independent coordinates in the equilibrium of Gordon-Newell networks (see Theorem A.2.6).

It should be noted that the three-term structure of the upper stratum of the product form steady state π in Theorem 3.4.4 can not be obtained from the general theory about "Jackson network in a random environment".

4. Inventory systems with perishable items

One of the key assumptions of inventory models is that the items in the inventory are always available for satisfying demand at any time. However, in certain types of inventories the items either perish, deteriorate, are subject to ageing, or become obsolete. For example products like foodstuffs, human blood, chemicals, etc. have a maximum life time.

There are many papers describing mathematical models of inventory systems, but a relatively small number consider items with a finite life time (cf. [Nah11, p. 1]). Analysis of inventory systems with perishable items is far more difficult than their counterpart having infinite life time. This is primarily because the inventory depletion rate is a function of the on-hand inventory level (cf. [ARS15, p. 26]). Despite their complexity, perishable inventory systems are investigated in many papers because of its potential applications in sectors like food, chemicals, pharmaceuticals, photography and blood bank management (cf. [ARS15, p. 26], [KSWD11, p. 393], [LC97, p. 1022]).

This chapter is devoted to perishable items and represents an extension of the basic model, which has been introduced in Chapter 2. The systems under investigation differ with respect to the distribution of the life time of raw material: Exponential distribution in Section 4.2.3 and general phase-type distribution in Section 4.2.2.

4.1. Related literature and own contributions

We point only the most relevant literature on continuous review perishable inventory models for our present investigation.

Literature reviews:
We refer to some articles which present a *literature review* of models for perishable inventory systems.

A review on the early literature on perishable inventory systems is provided by Nahmias [Nah82]. Raafat [Raa91] provides a survey of published literature for continuously deteriorating inventory models, where the deterioration is a function of the on-hand inventory. Shah and Shah [SS00] make a literature survey on inventory models for deteriorating items. Goyal and Giri [GG01] give a comprehensive literature review of models for inventory control of perishable items since the early 1990s. Bakker and his coauthors [BRT12] make an extensive review of more recent literature that has been published since the review of Goyal and Giri. In the article [JCS16] of Janssen and her coauthors, the literature reviews on inventory models for perishable goods are depicted graphically in [JCS16, Figure 1, p. 87]. They give an up-to-date review of perishable inventory models as well as of the joint key topics of publications from January 2012 until December 2015 in the research area of inventory models with perishable items.

Nahmias [Nah10] provides an overview of the basic theory and methods for modelling perishable inventory problems. His article focuses more on periodic review of a single location inventory system and it is a more technical discussion of the techniques employed in the analysis of perishable items.

Baron [Bar10] discusses inventory management of perishable items, including the differences between inventory models that consider items' perishability and the standard models that ignore this issue. The most common models for perishability are (1) outdatedness due to reaching expiry date (e.g. food items or medicine), which is typically modelled as a deterministic time and (2) sudden perishability due to disaster (e.g. spoilage because of extreme weather conditions), which is typically modelled as an exponential (or its discrete counterpart, geometric) time.

For recent overviews of the literature on perishable inventory systems we refer to Karaesmen, Scheller-Wolf and Deniz [KSWD11]. They provide an overview of research in Section 15.3.2 on continuous review single-location inventory models with a finite life time. Moreover, [Nah11] gives an extensive review of the inventory management of perishable products. Furthermore, Bijvank and Vis [BV11, Section 7.5, p. 10] and Krishnamoorthy et al. [KLM11] present a review of inventory models including perishable inventory systems.

Literature on perishable inventory systems with base stock policy:
Focus of our present investigations is the base stock policy. Thus, we point out the most relevant sources which deal with this replenishment policy and perishable items.

The paper of Schmidt and Nahmias [SN85] is the first that investigate analytically a model with positive lead times. They analyse a continuous review base stock policy for a perishable inventory system with fixed life and lead times under Poisson demand and lost sales. A generalization of this paper is the work of Perry and Posner [PP98] by considering a waiting time policy.

Pal [Pal89] and Liu and Cheung [LC97] consider a continuous review base stock policy for a perishable inventory system with exponentially distributed life and lead times. Similar to our approach they use methods and models from queueing theory to evaluate the performance of base stock policies.

Liu and Cheung consider two replenishment mechanisms: (1) each replenishment order is processed immediately regardless of the number of existing outstanding orders; (2) replenishment orders never cross and are processed one at a time according to a FCFS criteria (i.e. the lead time of an order depends on the number of outstanding orders). In [LC97] the possibilities of partial backorders, complete backorders and complete lost sales are all included. For these systems they derive the steady state probabilities by solving the global balance equations and develop some performance measures to perform a cost analysis.

In [Pal89] backorders are allowed. He derives the steady state probabilities by solving the global balance equations and develop some performance measures to perform a cost analysis.

Literature on production-to-order models with perishable inventory systems:
In most of the models considered in the literature, there is no production-to-order such that the time to satisfy customer demand is zero. In the following, we present some articles which consider continuous review models where the demanded items are issued to the customer only after some service is performed.

Manuel, Sivakumar, Arivarignan [MSA07] study an (r, Q)-policy for a perishable inventory system attached to a service facility with a finite waiting room and a single server, two types of customers and they adopt a removal rule. The service time is phase-type distributed and the life times and lead times are exponentially distributed. They obtain the joint probability distribution of the number of customers in the system and the inventory level in steady state by a recursive computation. Moreover, they compute various performance measures and calculate the total expected cost rate.

Manuel, Sivakumar and Arivarignan [MSA08] analyse an (r, Q)-policy for a perishable inventory system attached to a service facility with a finite waiting room and a single server with phase-type distribution and exponentially distributed life times, lead times and retrial times of orbiting customers. They derive the joint probability distribution of the number of customers in the waiting room, number of customers in the orbit and the inventory level in steady state by using matrix geometric methods (recursive computation) and obtain a stability condition. Furthermore, they compute various performance measures and calculate the total expected cost rate.

Krishnamoorthy and Anbazhagan [KA08] investigate an (r, Q)-policy for a perishable inventory system attached to a service facility with finite waiting room. They consider the N-policy which means that the service starts only if the customer level reaches N. They derive the joint stationary distribution of the number of customers in the system and the inventory levels. Furthermore, they develop various performance measures and illustrate the results with numerical examples.

Jeganathan [Jeg14] studies a $(0, Q)$-policy for a perishable inventory system attached to a service facility with multiple server vacations with a finite waiting room, impatient customers and zero lead time. The joint stationary distribution of the number of customers in the system and the inventory levels is obtained algorithmically. He derives various performance measures and calculates the total expected cost rate.

Amirthakodi, Radhamani and Sivakumar [ARS15] investigate a perishable inventory system attached to a service facility with a finite waiting room and a single server with positive service times and lead times and with feedback customers. The inventory is controlled by a variable ordering policy. They derive the steady state distribution of the system using matrix recursive methods and obtain a stability condition. Furthermore, they derive various performance measures, calculate the total expected cost rate and illustrate the results numerically.

Yadavalli, Anbazhagan and Jeganathan [YAJ15b] consider an (r, Q)-policy for a perishable inventory systems attached to a service facility having two heterogeneous exponential servers with one unreliable server and repeated attempts. The demands originate from a finite population. They derive the joint stationary distribution of the number of customers in the orbit and the inventory level. They develop various performance measures, calculate the total expected cost rate and illustrate the results numerically.

Shajin and her coauthors [SBDK16] consider a queueing-inventory model with reservation, cancellation. The items in a batch have a common life time. They investigate the cases of zero and of positive lead time and derive the stationary distribution of the queue length and the on-hand inventory in product form. Moreover, they study several performance measures.

Connection to Literature on impatient customers:
Perishable items of raw material arriving at an inventory can be treated as impatient customers arriving at a service station (cf. [Kum16, p. 98]), so that the shipment of these items from the inventory to the production system is a service provided to the waiting customers in the system.

A review on queueing systems for impatient customers is presented by Wang et al. [WLJ10]. A recent article about queueing-inventory with impatient customers is [MPB16].

Our main contributions are the following:
In the first part of this chapter, we study single location production-inventory models with queue-length-dependent arrival and service rates. We develop Markov process models for a supply chain which consists of a production system with an attached inventory with perishable raw material, where the item of raw material that is in the production process cannot perish. The systems under investigation differ with respect to the distribution of the life time of the raw material.

For the model with exponentially distributed life time we derive a sufficient condition and a necessary condition of ergodicity and prove that the stationary distribution has no product form. For the special case with base stock level equal to one, we obtain an explicit closed solution for the stationary distribution. For the system with base stock level greater than one we obtain some properties of the stationary system which provide insights into the equilibrium behaviour of the systems but an explicit expression of the complete stationary distribution is still an open problem.

For the model with phase-type distributed life time we derive a sufficient and necessary condition of ergodicity. We prove that the stationary distribution has no product form and we obtain some properties of the stationary system which provide insights into the equilibrium behaviour of the systems but an explicit expression of the complete stationary distribution is still an open problem.

Our work is an extension of the investigations of Pal [Pal89] and Liu and Cheung [LC97] on perishable inventory systems under base stock policy: In their models there is no production-to-order such that the time to satisfy customer demand is zero. Therefore, their model is a special case of our model with exponentially distributed life time when the service time is set to zero. We need to mention that Pal considers the backordering case and Liu and Cheung allow order crossing and also investigate the case of partial or complete backordering.

In the second part of this chapter, we modify the system with exponentially distributed life times so that we get product form results. More precisely, we take not into consideration whether there are customers in the system. The product form result is even true for a supply chain with $J > 1$ locations.

Furthermore, we deal with the question "Can we use the product form results from Section 4.3.2 to obtain simple product form bounds for the system with unknown non-product form stationary distribution in Section 4.2.3?".

Our conjecture that we can find upper and lower product form bounds for the throughputs is supported by our results for a system with base stock level $b = 1$. Under additional conditions for the system's parameters we can tackle even the case of $b \geq 2$.

The articles in the literature about perishable inventory models with production-to-order are only single location models. Our results extend their settings to a multi-dimensional system, whereby we consider the base stock policy.

A cost analysis can be performed as for the basic model in Section 2.5 on page 29.

4.2. Non-separable systems: Single location

4.2.1. Description of the general model

The supply chain of interest is depicted in Figure 4.2.1. The location consists of a production system with an attached inventory with perishable raw material. The inventory is replenished by a supplier which manufactures raw material for the location.

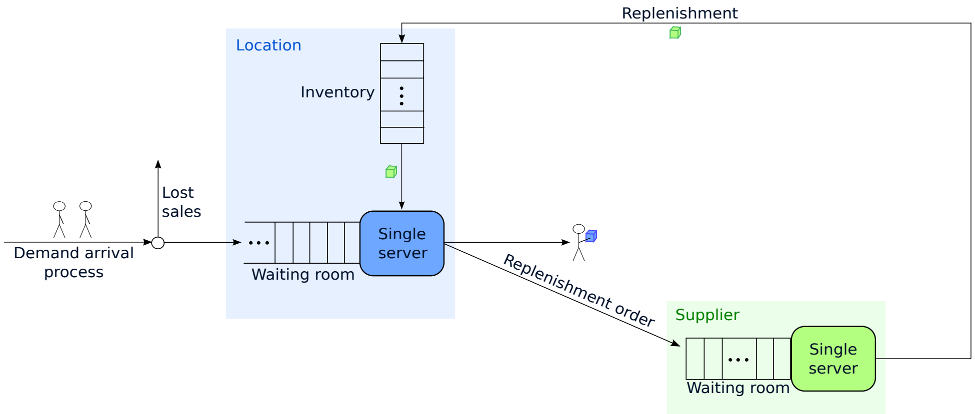


Figure 4.2.1.: Supply chain with base stock policy

Facilities in the supply chain. The production system consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under a FCFS regime. Customers arrive one by one at the production system according to a Poisson process with queue-length-dependent intensities and require service. To satisfy a customer's demand the production system needs exactly one item of raw material, which is taken from the associated local inventory. When a new customer arrives at the location while the previous customer's order is not finished, this customer will wait. If the inventory is depleted, the customers who are already waiting in line will wait, but new arriving customers at this location will decide not to join the queue and are lost ("local lost sales"). If there are $n \geq 0$ customers present and if the inventory is not depleted, customers arrive with intensity $\lambda(n) > 0$ at the production system.

The service requests at the location is exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at the location is provided with queue-length-dependent intensity. If there are $n > 0$ customers present either waiting or in service (if any) and if the inventory is not depleted, the service intensity is $\mu(n) > 0$. If the server is ready to serve a customer, who is at the head of the line, and the inventory is not depleted, the service immediately starts. Otherwise, the service starts at the instant of time when the next replenishment arrives at the local inventory.

For the control of the inventory we use the base stock policy. This means, each item taken from the inventory results in a direct order for one item of raw material. Hence, if a served customer departs from the system or a raw material is perished, an order of the consumed resp. of the perished raw material is placed at the supplier at this time instant. The base stock level $b \geq 1$ is the maximal size of the inventory. Note that there can be more than one outstanding order.

The items of raw material in the inventory are perishable. In the literature about perishable queueing-inventory systems (e.g. [MSA07, MSA08], [MSA08], [Jeg14], [YAJ15b]), it is often assumed that an item of raw material, which is in the production process cannot perish.

The systems under investigation differ with respect to the distribution of the life time of the raw material in the following way:

- phase-type distribution (Section 4.2.2),
- exponential distribution (Section 4.2.3), which is a special case of the phase-type distribution, but deserves interest because of simplicity of the results.

The supplier consists of a single server (machine) and a waiting room under FCFS regime. At most $b - 1$ replenishment orders are waiting at the supplier. Service times at the supplier are exponentially distributed with parameter $\nu > 0$.

Routing in the supply chain. A customer departs from the system immediately after the service and the associated consumed raw material is removed from the inventory at this time instant. A finished item of raw material departs immediately from the supplier and the item is added to the inventory at the location.

We assume that the replenished raw materials are "fresh". Much of the literature on perishable items assumes this to avoid to complicate the model (cf. [Bar10, p. 2]).

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the supplier and the inventory are negligible.

All inter-arrival times, service times and life times of items constitute an independent family of random variables.

Remark 4.2.1. These models are special cases of queueing systems in a random environment which we have introduced in Appendix D.1. In Example D.1.2 on page 326 we show the setting by which the queueing-inventory systems described in this chapter fit into the definition of the queueing systems in a random environment. An overview of the corresponding results is presented in Table D.1 on page 348. We provide some direct proofs here to avoid the general lengthy formalism needed in queueing systems in a random environment.

4.2.2. Phase-type distributed life time

This section is dedicated to the study of the supply chain with one location ($J = 1$, $\overline{J} = \{1\}$) as described in Section 4.2.1, where the life time of each raw material is phase-type distributed with mean γ^{-1} and the item of raw material that is in the production process cannot perish.

Remark 4.2.2. The most common model in the literature to count the inventory is the following (cf. Section 4.1): One item is removed from the inventory when a served customer departs from the system. In the present model with phase-type distributed life time of perishable items, it is important to know whether a specific item is in production or on stock, because it is assumed (e.g. [MSA07, MSA08], [MSA08], [Jeg14], [YAJ15b]) that an item which is in production cannot perish (see Section 4.2.1).

It should be noted that in all other sections we do not need this essential distinction.

We consider life time distributions of the following phase-type which are sufficient versatile to approximate any distribution on $\mathbb{R}_0^+$ arbitrary close. The next definition is based on [Dad01b, Definition 9.2 , pp. 347f.].

Definition 4.2.3. For $h \in \mathbb{N}$ and $\beta > 0$ let

$$\Gamma_{\beta,h}(s) = 1 - e^{-\beta s} \sum_{i=0}^{h-1} \frac{(\beta s)^i}{i!}, \qquad s \geq 0,$$

denote the cumulative distribution function of the Γ-distribution with parameters β and h. The parameter h is a positive integer and serves as a phase-parameter for the number of independent exponential phases, each with mean β^{-1}, the sum of which constitutes a random variable with distribution $\Gamma_{\beta,h}$. ($\Gamma_{\beta,h}$ is called a h-stage Erlang distribution with scale parameter β.)

We consider the following class of distributions on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$, which is dense with respect to the topology of weak convergence of probability measures in the set of all distributions on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$ [Sch73, Section I.6]. For $\beta \in (0, \infty)$, $H \in \mathbb{N}$, and probability $b(\cdot)$ on $\{1, \ldots, H\}$ with $b(H) > 0$ let the cumulative distribution function

$$B(s) = \sum_{\ell=1}^{H} b(\ell) \cdot \Gamma_{\beta,\ell}(s), \qquad s \geq 0,$$

denote a phase-type distribution function. With varying β, H and $b(\cdot)$ we can approximate any distribution on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$ sufficiently close.

We refer to [Asm03, Chapter III.4] for a short introduction into this and various other classes of phase-type distributions.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X(t)$ the number of customers present at the location at time $t \geq 0$, either waiting or in service (queue length). By $Y(t)$ we denote the state of the inventory, whereby we have to distinguish whether an item of raw material is in production or on stock, at time $t \geq 0$. By $W(t)$ we denote the number of replenishment orders at the central supplier at time $t \geq 0$, either waiting or in service (queue length).

We define the joint queueing-inventory process of this system by

$$Z = ((X(t), Y(t), W(t)) : t \geq 0) .$$

To describe the state of the inventory, we have to distinguish whether an item of raw material is in production or on stock. For the items of raw material on stock we need the number of residual phases of each item in the inventory. However, we do not need the information about the phase of the item in production, since the item of raw material, which is in the production process cannot perish. This leads to

$$\mathbf{k} = \Big(\underbrace{\overbrace{\overbrace{\overline{k}}^{\substack{\text{number of items}\\ \text{in production}}} \quad , \quad \overbrace{\underbrace{[h_1, \dots, h_k]}_{\substack{k=\text{number of items}\\ \text{on stock}}}}^{\substack{\text{number of residual phases}\\ \text{of lifetimes}}}}^{\text{inventory}}}_{\substack{\overline{k}+k=\text{number of items}\\ \text{in inventory}}} , \overbrace{b - (\overline{k} + k)}^{\text{supplier}} \Big).$$

- $\overline{k} \in \{0, 1\}$ indicates whether there is an item in production or not,
- k denotes the number of items on stock with $k \in \{0, \dots, b\}$ if $\overline{k} = 0$ and $k \in \{0, \dots, b-1\}$ if $\overline{k} = 1$,
- $\overline{k} + k$ is the total number of items at the location,
- $h_j \in \{1, \dots, H\}$ is the number of residual phases of the item on position j, $0 \leq j \leq k$, with $h_1 \leq h_2 \leq \dots \leq h_k$, $0 \leq k + \overline{k} \leq b$, i.e. items are sorted in line by their phases in ascending order. This means if $k > 0$ items are on stock, that on position 1 resides an item with the smallest number of phases and on position k resides an item with the highest number of phases.
  Note that position numbers are associated only with items on stock (waiting). To items in production no position number is associated.

This results in

$$\begin{aligned} K := \; & \Big\{ \Big(\overline{k}, [h_1, \dots, h_k], b - (\overline{k} + k)\Big) : \\ & \quad \overline{k} \in \{0,1\}, h_\ell \in \{1, \dots, H\},\ 1 \leq \ell \leq k,\ h_1 \leq h_2 \leq \dots \leq h_k,\ 0 \leq \overline{k} + k \leq b \Big\} \\ = \; & \Big\{ (0, [h_1, \dots, h_k], b - k) : \\ & \quad h_\ell \in \{1, \dots, H\},\ 1 \leq \ell \leq k,\ h_1 \leq h_2 \leq \dots \leq h_k,\ 0 \leq k \leq b \Big\} \\ & \cup \Big\{ (1, [h_1, \dots, h_k], b - (1 + k)) : \\ & \quad h_\ell \in \{1, \dots, H\},\ 1 \leq \ell \leq k,\ h_1 \leq h_2 \leq \dots \leq h_k,\ 0 \leq k \leq b - 1 \Big\}. \end{aligned}$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process. The state space of Z is

$$E = \{(n, \mathbf{k}) : n \in \mathbb{N}_0,\, \mathbf{k} \in K\},$$

where n is the number of customers and $\mathbf{k}$ describes the state of the inventory and the central supplier.

Note the redundancy in the state space: $W(t) = b - Y(t)$. We prefer to carry all information explicitly with because the dynamics of the system are easier visible.

For example,
if $\ell > 1$ items of raw material are in the inventory and there is at least one customer, then

$$(n,\mathbf{k}) = \Bigg(n, \underbrace{\overbrace{\overbrace{1}^{\substack{\text{number of items}\\ \text{in production}}}\;,\; \overbrace{\underbrace{[h_1,\dots,h_{\ell-1}]}_{\substack{\ell-1=\text{number of items}\\ \text{on stock}}}}^{\substack{\text{number of residual phases}\\ \text{of life times}}}}^{\text{inventory}}}_{\substack{\ell=\text{number of items}\\ \text{in inventory}}}, \overbrace{b-\ell}^{\text{supplier}}\Bigg),$$

if $\ell > 1$ items of raw material are in the inventory and there is no customer, then

$$(n,\mathbf{k}) = \Bigg(0, \underbrace{\overbrace{\overbrace{0}^{\substack{\text{number of items}\\ \text{in production}}}\;,\; \overbrace{\underbrace{[h_1,\dots,h_{\ell}]}_{\substack{\ell=\text{number of items}\\ \text{on stock}}}}^{\substack{\text{number of residual phases}\\ \text{of life times}}}}^{\text{inventory}}}_{\substack{\ell=\text{number of items}\\ \text{in inventory}}}, \overbrace{b-\ell}^{\text{supplier}}\Bigg),$$

if $\ell = 1$ items of raw material are in the inventory and there is at least one customer, then

$$(n,\mathbf{k}) = \Bigg(n, \underbrace{\overbrace{\overbrace{1}^{\substack{\text{number of items}\\ \text{in production}}}\;,\; \overbrace{[0]}^{\substack{\text{number of residual phases}\\ \text{of life times}}}}^{\text{inventory}}}_{\substack{1=\text{number of items}\\ \text{in inventory}}}, \overbrace{b-1}^{\text{supplier}}\Bigg),$$

if $\ell = 1$ items of raw material are in the inventory and there is no customer, then

$$(n,\mathbf{k}) = \Bigg(0, \underbrace{\overbrace{\overbrace{0}^{\substack{\text{number of items}\\ \text{in production}}}\;,\; \overbrace{\underbrace{[h_1]}_{\substack{1=\text{number of items}\\ \text{on stock}}}}^{\substack{\text{number of residual phases}\\ \text{of life times}}}}^{\text{inventory}}}_{\substack{1=\text{number of items}\\ \text{in inventory}}}, \overbrace{b-1}^{\text{supplier}}\Bigg).$$

This page

intentionally

left blank.

4.2.2.1. Ergodicity

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z; z) : z, \tilde{z} \in E)$ with the following transition rates for $(n, \mathbf{k}) \in E$, where a typical state is

$$(n,\mathbf{k}) = \Bigg(n, \underbrace{\overbrace{\overbrace{\overline{k}}^{\substack{\text{number of items}\\ \text{in production}}} \quad , \quad \underbrace{\overbrace{[h_1,\dots,h_k]}^{\substack{\text{number of residual phases}\\ \text{of life times}}}}_{\substack{k=\text{number of items}\\ \text{on stock}}}}^{\text{inventory}}}_{\substack{\overline{k}+k=\text{number of items}\\ \text{in inventory}}}, \overbrace{b-(\overline{k}+k)}^{\text{supplier}}\Bigg)$$

and we will impose necessary restrictions if needed:[1]

- Arrival of a customer,
  which happens only if there is at least one item at the location ($\overline{k} + k > 0$), either on stock or in production, because of the lost sales rule:
  $$q\left((n,\mathbf{k});(n+1,\mathbf{k})\right) = \lambda(n) \cdot 1_{\{\overline{k}+k>0\}}, \qquad n \geq 0.$$

- Service completion of a customer,
  which happens only if there is at least one customer ($n > 0$) and an item in production ($\overline{k} = 1$), i.e. a customer departs from the location and takes an item.
  Let $T_{0\searrow}\mathbf{k}$ be the state after the following event:
  The item, which is in production, is removed from the location and a replenishment order is sent to the central supplier, where $b - (\overline{k} + k)$ orders have already been present.
  If $n - 1 > 0$, the item on position 1 is taken in production and the items previously on positions $2, \dots, k$ move to positions $1, \dots, k - 1$
  and if $n - 1 = 0$, there are no changes of the positions, because there are no items on stock:
  $$q\Big(\big(n,\mathbf{k}\big);\big(n-1, T_{0\searrow}\mathbf{k}\big)\Big) = \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{\overline{k}=1\}}.$$

[1] Position numbers are associated only with items on stock (waiting). To items in production no position number is associated.

- PHASE COMPLETION OF A PERISHABLE ITEM ON POSITION ℓ, $\ell \in \{1,\dots,k\}$, which happens only if there is at least one item on stock ($k > 0$), i.e. either

- if $h_\ell = 1$, i.e. there is a perishable item on position ℓ, which is in its last phase. Let $T_{\ell\searrow}\mathbf{k}$ be the state after the following event (see Figure 4.2.2(a)):
This perishable item on position ℓ is removed from the location and an order of one unit is sent to the central supplier, where $b - (\overline{k} + k)$ orders have already been present, and the perishable items previously on positions $\ell+1, \ell+2, \dots, k$ move to positions $\ell, \ell+1, \dots k-1$:

$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\ell\searrow}\mathbf{k}\big)\Big) = \beta \cdot 1_{\{h_\ell=1\}} \cdot 1_{\{k>0\}}, \quad n \geq 0.$$

Note, that if $k = 1$, the list of positions, where other items sit, is empty. Consequently, there are no further changes in the positions. More precisely, for $k = 1$ holds

$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\ell\searrow}\mathbf{k}\big)\Big) = q\Big(\big(n,\overline{k},[1],b-(\overline{k}+1)\big);\big(n,\overline{k},[0],b-(\overline{k}+1)+1\big)\Big).$$

- or if $h_\ell > 1$, i.e. there is a perishable item on position ℓ, which is not in its last phase, and either

  ⋆ if $k > 0$ and $h_\ell - 1 < h_1$,
  let $T_{\ell\to 1}\mathbf{k}$ be the state after this event (see Figure 4.2.2(b)):
  The phase of this perishable item on position ℓ is shifted one step down and the item is moved to position 1 and the items on positions $1,\dots,\ell-1$ move to positions $2,\dots,\ell$:

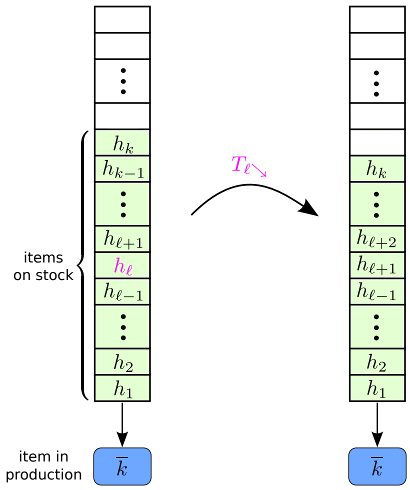
$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\ell\to 1}\mathbf{k}\big)\Big) = \beta \cdot 1_{\{0<h_\ell-1<h_1\}} \cdot 1_{\{k>0\}}, \quad n \geq 0.$$

  Note, that if $\ell = 1$, i.e. the phase of the item on position 1 is shifted one step down and there are no changes in the positions.
  Furthermore, if $k = 1$, the list of positions, where other items sit, is empty. Consequently, there are no further changes in the positions.

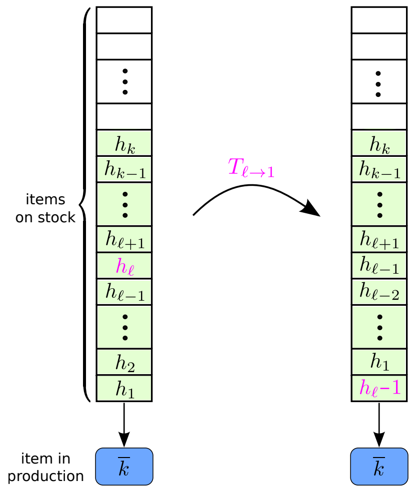
  ⋆ if $k > 1$ and $h_{m-1} \leq h_\ell - 1 < h_m$, $m \in \{2,\dots,\ell\}$,
  let $T_{\ell\to m}\mathbf{k}$ be the state after this event (see Figure 4.2.2(c)):
  The phase of this perishable item on position ℓ is shifted one step down and the item is moved to position m, $m \in \{2,\dots,\ell\}$, and the items on positions $m,\dots,\ell-1$ move to positions $m+1,\dots,\ell$:

$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\ell\to m}\mathbf{k}\big)\Big) = \beta \cdot 1_{\{0<h_{m-1}\leq h_\ell-1<h_m\}} \cdot 1_{\{k>1\}} \cdot 1_{\{m\in\{2,\dots,\ell\}\}}, \; n \geq 0.$$

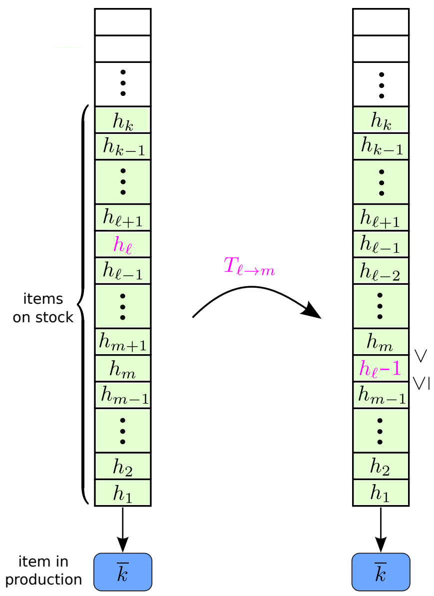
  Note, that if $m = \ell$, the list of positions where other items sit is empty. Consequently, there are no further changes in the positions.



(a) $k > 0$, $h_\ell = 1$



(b) $k > 0$, $0 < h_\ell - 1 < h_1$



(c) $k > 1$, $h_{m-1} \leq h_\ell - 1 < h_m$, $m \in \{2, \ldots, \ell\}$

Figure 4.2.2.: Changes in $\mathbf{k}$ after a phase completion of a perishable item on position ℓ, $\ell \in \{1, \ldots, k\}$

- Service completion of an order at the central supplier, which happens only if at the central supplier is at least one order ($b-(\overline{k}+k)>0$), i.e. an item is removed from the central supplier and is sent to the location and if appropriate, it becomes $\widetilde{h}$ phases with probability $b(\widetilde{h})$, and either

- if $\overline{k}+k=0$, i.e. there are no items in the inventory, and either
  - ⋆ if $n=0$, i.e. there are no customers, then the item is placed on position 1 of the inventory and obtains $\widetilde{h}$ phases with probability $b(\widetilde{h})$:

$$q\Big(\big(0,0,[0],b\big);\big(0,0,[\widetilde{h}],b-1\big)\Big)=\nu\cdot b(\widetilde{h}),$$

  - ⋆ or if $n>0$, i.e. there is at least one customer, then the item is taken immediately in production:

$$q\Big(\big(n,0,[0],b\big);\big(n,1,[0],b-1\big)\Big)=\nu\cdot 1_{\{n>0\}}.$$

- or if $\overline{k}+k>0$, i.e. there is at least one item in the inventory, then the item obtains $\widetilde{h}$ phases with probability $b(\widetilde{h})$ and either
  - ⋆ if $k=0$ and $\overline{k}=1$, the item moves into position 1 (note that there must be at least one customer ($n>0$) because there is an item in production ($\overline{k}=1$)):

$$q\Big(\big(n,1,[0],b\big);\big(n,1,[\widetilde{h}],b-1\big)\Big)=\nu\cdot b(\widetilde{h})\cdot 1_{\underbrace{\{0<\overline{k}+k<b\}}_{=\{b>1\}}},\ n>0,$$

  - ⋆ or if $k>0$, $\overline{k}\in\{0,1\}$ and $\widetilde{h}<h_1$, the item moves into position 1. The items previously on positions $1,2,\dots,k$ move to positions $2,3,\dots,k+1$. Let $T_{\widetilde{h}\to 1}\mathbf{k}$ be the state after this event (see Figure 4.2.3(a)):

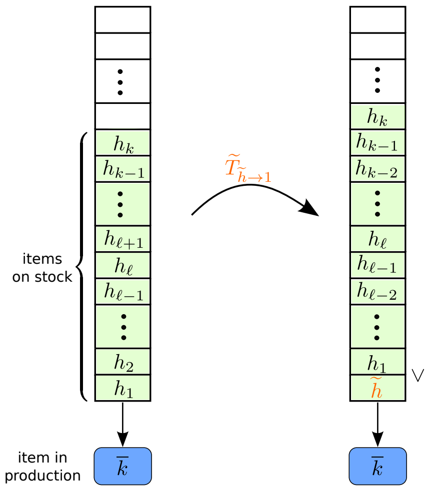
$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\widetilde{h}\to 1}\mathbf{k}\big)\Big)=\nu\cdot b(\widetilde{h})\cdot 1_{\{\widetilde{h}<h_1\}}\cdot 1_{\{0<\overline{k}+k<b\}}\cdot 1_{\{k>0\}},\quad n\geq 0,$$

  - ⋆ or if $k>0$, $\overline{k}\in\{0,1\}$ and $h_k\leq\widetilde{h}$, the item moves into position $k+1$. Let $T_{\widetilde{h}\to k+1}\mathbf{k}$ be the state after this event (see Figure 4.2.3(b)):

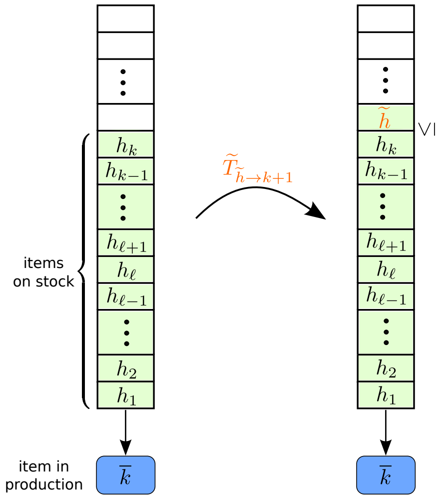
$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\widetilde{h}\to k+1}\mathbf{k}\big)\Big)=\nu\cdot b(\widetilde{h})\cdot 1_{\{h_k\leq\widetilde{h}\}}\cdot 1_{\{0<\overline{k}+k<b\}}\cdot 1_{\{k>0\}},\quad n\geq 0,$$

  - ⋆ or if $k>1$, $\overline{k}\in\{0,1\}$ and $h_{\ell-1}\leq\widetilde{h}<h_\ell$, $\ell\in\{2,\dots k\}$, the item moves into position ℓ. The items previously on positions $\ell,\ell+1,\dots,k$ move to positions $\ell+1,\ell+2,\dots,k+1$. Let $T_{\widetilde{h}\to\ell}\mathbf{k}$ be the state after this event (see Figure 4.2.3(c)):

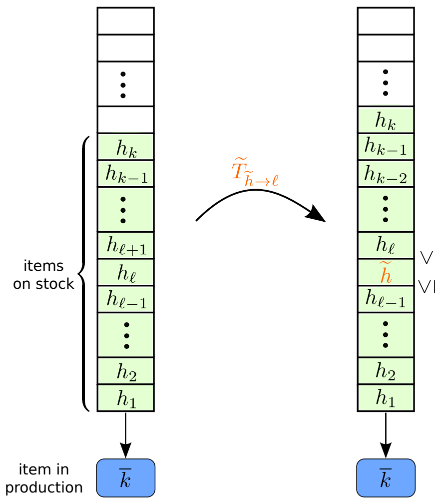
$$q\Big(\big(n,\mathbf{k}\big);\big(n,T_{\widetilde{h}\to\ell}\mathbf{k}\big)\Big)$$
$$=\nu\cdot b(\widetilde{h})\cdot 1_{\{h_{\ell-1}\leq\widetilde{h}<h_\ell\}}\cdot 1_{\{0<\overline{k}+k<b\}}\cdot 1_{\{k>1\}}\cdot 1_{\{\ell\in\{2,\dots,k\}\}},\quad n\geq 0,$$



(a) $0 < \overline{k} + k < b,\ k > 0,\ \widetilde{h} < h_1$



(b) $0 < \overline{k} + k < b,\ k > 0,\ h_k \leq \widetilde{h}$



(c) $0 < \overline{k} + k < b,\ k > 1,\ h_{\ell-1} \leq \widetilde{h} < h_\ell,$
$\ell \in \{2, \ldots, k\}$

Figure 4.2.3.: Changes in $\mathbf{k}$ after service completion of an order at the central supplier

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

From the above transition rates follows similarly as in Appendix E.1 that Z is irreducible on the state space E.

We will show a necessary condition for ergodicity in Proposition 4.2.5. Furthermore, a sufficient condition for positive recurrence is shown in Proposition 4.2.6.

The proof for the necessary condition is a special case of Proposition D.1.4 (and Proposition D.1.3). We provide a direct proof here to avoid the general lengthy formalism needed in queueing-environment systems.

The proof for the necessary condition depends on the following result.

Proposition 4.2.4. *If the queueing-inventory process Z is recurrent, then any solution $\mathbf{x} = (x(z) : z \in E)$ of the global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ fulfils for all $n \in \mathbb{N}_0$*

$$\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n,\mathbf{k}) = \sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n+1,\mathbf{k}) \cdot \frac{\mu(n+1)}{\lambda(n)} \tag{4.2.1}$$

and

$$\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n,\mathbf{k}) = \sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(0,\mathbf{k}) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \tag{4.2.2}$$

Proof. From irreducibility and recurrence of Z follows that there exists one, and up to a multiplicative factor only one, stationary measure $\mathbf{x} = (x(z) : z \in E)$. This stationary measure $\mathbf{x}$ has the property $x(z) > 0$ for all $z \in E$ and can be found as a solution of the global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ (cf. [Asm03, Theorem 4.2, p. 51]).

Equation (4.2.1) can be proven by the cut-criterion for recurrent processes, which is presented in Theorem A.1.1(b). For $n \in \mathbb{N}_0$, equation (4.2.1) can be proven by a cut, which divides E into complementary sets according to the queue length of customers that is less than or equal to n or greater than n, i.e. into the sets

$$\Big\{(m,\mathbf{k}) : m \in \{0,1,\dots,n\},\ \mathbf{k}\in K\Big\},$$
$$\Big\{(\widetilde{m},\widetilde{\mathbf{k}}) : \widetilde{m} \in \mathbb{N}_0 \setminus \{0,1,\dots.n\},\ \widetilde{\mathbf{k}}\in K\Big\}, \quad n \in \mathbb{N}_0.$$

Thus, it follows for $n \in \mathbb{N}_0$

$$\sum_{m=0}^{n}\sum_{\mathbf{k}\in K}\sum_{\widetilde{m}=n+1}^{\infty}\sum_{\widetilde{\mathbf{k}}\in K} x(m,\mathbf{k})\cdot q((m,\mathbf{k});(\widetilde{m},\widetilde{\mathbf{k}}))$$
$$= \sum_{\widetilde{m}=n+1}^{\infty}\sum_{\widetilde{\mathbf{k}}\in K}\sum_{m=0}^{n}\sum_{\mathbf{k}\in K} x(\widetilde{m},\widetilde{\mathbf{k}})\cdot q((\widetilde{m},\widetilde{\mathbf{k}});(m,k))$$

$$\Leftrightarrow \qquad \sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n,\mathbf{k})\cdot\lambda(n) = \sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n+1,\mathbf{k})\cdot\mu(n+1).$$

Consequently, for $n \in \mathbb{N}_0$ follows equation (4.2.2)

$$\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n,\mathbf{k}) = \sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(0,\mathbf{k}) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}.$$

□

Proposition 4.2.5. *If queueing-inventory process Z is ergodic, it holds $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} < \infty$.*

Proof. If the queueing-inventory process Z is ergodic, the normalisation constant, as the sum of the solution $\mathbf{x} = (x(z) : z \in E)$ of the global balance equation $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ for Z, is finite,

$$\sum_{n=0}^{\infty}\sum_{\mathbf{k}\in K} x(n,k,b-k) < \infty.$$

It holds

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{\mathbf{k}\in K} x(n,\mathbf{k}) \\
&= \sum_{n=0}^{\infty}\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(n,\mathbf{k}) + \sum_{n=0}^{\infty} x(n,0,[0],b) \\
&\overset{(4.2.2)}{=} \sum_{n=0}^{\infty}\underbrace{\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} x(0,\mathbf{k})}_{=:\widetilde{W}} \cdot \prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} + \sum_{n=0}^{\infty} x(n,0,[0],b) \\
&= \widetilde{W}\cdot\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} + \sum_{n=0}^{\infty} x(n,0,[0],b).
\end{aligned}$$

Because of ergodicity, $\widetilde{W} \in (0,\infty)$ and $\sum_{n=0}^{\infty} x(n,0,[0],b) < \infty$.
Hence, $\sum_{n=0}^{\infty}\sum_{\mathbf{k}\in K} x(n,\mathbf{k})$ is finite only if $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} < \infty$. □

Proposition 4.2.6.

(a) *The queueing-inventory process Z is ergodic if for an $M/M/1/\infty$ queue with queue-length dependent arrival intensities $\lambda(n) > 0$ and service intensities $\mu(n) > 0$ there exists a Lyapunov function $\widetilde{\mathcal{L}} : \mathbb{N}_0 \to \mathbb{R}_0^+$ with finite exception set $\widetilde{F}$ and constant $\widetilde{\varepsilon} > 0$, which satisfies the Foster-Lyapunov stability criterion, and* $\sup_{n\in\mathbb{N}_0}\mu(n) < \infty$.

(b) *The queueing-inventory process Z is ergodic if there exists $N \in \mathbb{N}_0$ such that $\inf_{n\geq N}(\mu(n)-\lambda(n)) > 0$ and* $\sup_{n\in\mathbb{N}_0}\mu(n) < \infty$.

Remark 4.2.7. The condition $\sup_{n\in\mathbb{N}_0}\mu(n) < \infty$ can be weakened by $\inf_{n\in\mathbb{N}_0}\widehat{c}_n > 0$, where $\widehat{c}_n$ is defined in Lemma D.1.7.

Proof of Proposition 4.2.6. As mentioned before, the queueing-inventory model is a special case of the queueing system in a random environment which we have introduced in Appendix D.1. Hence, sufficiency follows in **(a)** from Proposition D.1.8 and in **(b)** from Corollary D.1.9 since $\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0$ holds by Example D.1.11(b). □

Remark 4.2.8. The queueing-inventory system can be modelled as a level-dependent quasi-birth-and-death process (LDQBD process). Under the assumptions from the above proposition, the queueing-inventory system is ergodic and hence, we can use the algorithm of Bright and Taylor [BT95] to calculate the equilibrium distributions in LDQBD processes.

Special case: Queue-length-independent arrival and service rates

In this paragraph, we analyse the queueing-inventory system with state-independent service rates μ and arrival rates λ in view of a sufficient and necessary condition for ergodicity.

Theorem 4.2.9. *The queueing-inventory process Z is ergodic if and only if $\lambda < \mu$.*

Proof. As mentioned before, the queueing-inventory model is a special case of the queueing system in a random environment which we have introduced in Appendix D.1. Hence, the sufficient and necessary condition follows from Proposition D.1.12 and Proposition D.1.13 since $\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0$ holds by Example D.1.11(b). Additionally, we provide a direct proof here to avoid the general lengthy formalism needed in queueing-environment systems.

Necessity follows from Proposition 4.2.5 with $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} = \sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n < \infty$.

We will show sufficiency by using the Foster-Lyapunov stability criterion, which is presented in Theorem A.1.2 on page 260. We will show that $\mathcal{L} : E \to \mathbb{R}_0^+$ with

$$\mathcal{L}(n, \mathbf{k}) = n + 1_{\{\bar{k}+k=0\}} \cdot \frac{\mu - \lambda}{2\cdot\mu}$$

and the finite exception set $F = \{(n, \mathbf{k}) : n = 0\} \subsetneq E$ is a Lyapunov function.

We define

$$\varepsilon = \min\left(\frac{\nu}{\mu}\cdot\frac{\mu-\lambda}{2}, \frac{\mu-\lambda}{2}\right).$$

► First, we will check $(\mathbf{Q}\cdot\mathcal{L})(n, \mathbf{k}) < \infty$ for $(n, \mathbf{k}) \in F = \{(n, \mathbf{k}) : n = 0\}$.

Since $0 < \lambda < \mu < \infty$, $0 < \nu < \infty$ and $0 < \beta < \infty$,
for $(0, 0, [0], b)$ holds

$$(\mathbf{Q}\cdot\mathcal{L})(0, 0, [0], b) = \sum_{\widetilde{h}=1}^{H} \nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}(0, 0, [\widetilde{h}], b-1) - \mathcal{L}(0, 0, [0], b)\Big) < \infty,$$

if $b \geq 2$, for $(0,0,[h_1],b-1)$ holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})\,(0,0,[h_1],b-1)\\
=\;&\lambda\cdot\left(\mathcal{L}\,(1,0,[h_1],b-1)-\mathcal{L}\,(0,0,[h_1],b-1)\right)\cdot\underbrace{1_{\{\bar{k}+k>0\}}}_{=1}\\
&+\sum_{\widetilde{h}=1}^{h_1-1}\nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}\Big(0,0,[\widetilde{h},h_1],b-2\Big)-\mathcal{L}\,(0,0,[h_1],b-1)\Big)\cdot\underbrace{1_{\{\widetilde{h}<h_1\}}}_{=1}\cdot\underbrace{1_{\{0<\bar{k}+k<b\}}}_{=1}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\widetilde{h}=h_1}^{H}\nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}\Big(0,0,[h_1,\widetilde{h}],b-2\Big)-\mathcal{L}\,(0,0,[h_1],b-1)\Big)\cdot\underbrace{1_{\{h_1\leq\widetilde{h}\}}}_{=1}\cdot\underbrace{1_{\{0<\bar{k}+k<b\}}}_{=1}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\beta\cdot\left(\mathcal{L}\,(0,0,[h_1-1],b-1)-\mathcal{L}\,(0,0,[h_1],b-1)\right)\cdot 1_{\{h_1>1\}}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\beta\cdot\left(\mathcal{L}\,(0,0,[0],b)-\mathcal{L}\,(0,0,[h_1],b-1)\right)\cdot 1_{\{h_1=1\}}\cdot\underbrace{1_{\{k>0\}}}_{=1}<\infty,
\end{aligned}
$$

if $b > 2$, for $(0,\mathbf{k})=\Big(0,0,[h_1,\ldots,h_k]\,,b-k\Big)$ with $k=2,\ldots,b-1$ holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})\,(0,\mathbf{k})\\
=\;&\lambda\cdot\left(\mathcal{L}(1,\mathbf{k})-\mathcal{L}(0,\mathbf{k})\right)\cdot\underbrace{1_{\{\bar{k}+k>0\}}}_{=1}\\
&+\sum_{\widetilde{h}=1}^{h_1-1}\nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}(0,\widetilde{T}_{\widetilde{h}\to 1}\mathbf{k})-\mathcal{L}(0,\mathbf{k})\Big)\cdot\underbrace{1_{\{\widetilde{h}<h_1\}}}_{=1}\cdot\underbrace{1_{\{0<\bar{k}+k<b\}}}_{=1}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\widetilde{h}=h_{\ell-1}}^{h_\ell-1}\nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}(0,\widetilde{T}_{\widetilde{h}\to \ell}\mathbf{k})-\mathcal{L}(0,\mathbf{k})\Big)\cdot\underbrace{1_{\{h_{\ell-1}\leq\widetilde{h}<h_\ell\}}}_{=1}\cdot\underbrace{1_{\{0<\bar{k}+k<b\}}}_{=1}\cdot\underbrace{1_{\{k>1\}}}_{=1}\cdot 1_{\{\ell\in\{2,\ldots,k\}\}}\\
&+\sum_{\widetilde{h}=h_k}^{H}\nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}(0,\widetilde{T}_{\widetilde{h}\to k+1}\mathbf{k})-\mathcal{L}(0,\mathbf{k})\Big)\cdot\underbrace{1_{\{h_k\leq\widetilde{h}\}}}_{=1}\cdot\underbrace{1_{\{0<\bar{k}+k<b\}}}_{=1}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\ell=1}^{k}\beta\cdot\left(\mathcal{L}\,(0,T_{\ell\searrow}\mathbf{k})-\mathcal{L}\,(0,\mathbf{k})\right)\cdot 1_{\{h_\ell=1\}}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\ell=1}^{k}\beta\cdot\left(\mathcal{L}\,(0,T_{\ell\to 1}\mathbf{k})-\mathcal{L}\,(0,\mathbf{k})\right)\cdot 1_{\{0<h_\ell-1<h_1\}}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\ell=1}^{k}\beta\cdot\left(\mathcal{L}\,(0,T_{\ell\to m}\mathbf{k})-\mathcal{L}\,(0,\mathbf{k})\right)\cdot 1_{\{0<h_{m-1}\leq h_\ell-1<h_m\}}\cdot\underbrace{1_{\{k>1\}}}_{=1}\cdot 1_{\{m\in\{2,\ldots,\ell\}\}}<\infty,
\end{aligned}
$$

for $(0,\mathbf{k}) = \Big(0,0,[h_1,\dots,h_b],0\Big)$ with $k=b\geq 1$ holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})(0,\mathbf{k})\\
=\;&\lambda\cdot((1,\mathbf{k})-\mathcal{L}(0,\mathbf{k}))\cdot\underbrace{1_{\{\overline{k}+k>0\}}}_{=1}\\
&+\sum_{\ell=1}^{b}\beta\cdot(\mathcal{L}(0,T_{\ell\searrow}\mathbf{k})-\mathcal{L}(0,\mathbf{k}))\cdot 1_{\{h_\ell=1\}}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\ell=1}^{b}\beta\cdot(\mathcal{L}(0,T_{\ell\to 1}\mathbf{k})-\mathcal{L}(0,\mathbf{k}))\cdot 1_{\{0<h_\ell-1<h_1\}}\cdot\underbrace{1_{\{k>0\}}}_{=1}\\
&+\sum_{\ell=1}^{b}\beta\cdot(\mathcal{L}(0,T_{\ell\to m}\mathbf{k})-\mathcal{L}(0,\mathbf{k}))\cdot 1_{\{0<h_{m-1}\leq h_\ell-1<h_m\}}\cdot\underbrace{1_{\{b>1\}}}_{=1_{\{k>1\}}}\cdot 1_{\{m\in\{2,\dots,\ell\}\}}<\infty.
\end{aligned}
$$

► Second, we will check $(\mathbf{Q}\cdot\mathcal{L})(n,\mathbf{k})\leq-\varepsilon$ for $(n,\mathbf{k})\notin F$ with

$$
-\varepsilon=\max\left(\frac{\nu}{\mu}\cdot\frac{\lambda-\mu}{2}\;,\;\frac{\lambda-\mu}{2}\right)=\begin{cases}\frac{\lambda-\mu}{2}<0 & \text{if } \mu\leq\nu,\\ \frac{\nu}{\mu}\cdot\left(\frac{\lambda-\mu}{2}\right)<0 & \text{if } \mu>\nu.\end{cases}
$$

Note that $n>0$ and hence it holds $\overline{k}=1$ if $\overline{k}+k\geq 1$.
For $\overline{k}+k=0$ and $n>0$ holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})(n,0,[0],b)\\
=\;&\nu\cdot(\mathcal{L}(n,1,[0],b-1)-\mathcal{L}(n,0,[0],b))\cdot\underbrace{1_{\{n>0\}}}_{=1}\\
=\;&\nu\cdot\left(n-n-\frac{\mu-\lambda}{2\cdot\mu}\right)=\frac{\nu}{\mu}\cdot\frac{\lambda-\mu}{2}\leq-\varepsilon.
\end{aligned}
$$

If $b\geq 2$, for $(n,1,[0],b-1)$ with $\overline{k}+k=1$ and $n>0$ holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})(n,1,[0],b-1)\\
=\;&\lambda\cdot(\mathcal{L}(n+1,1,[0],b-1)-\mathcal{L}(n,1,[0],b-1))\cdot\underbrace{1_{\{\overline{k}+k>0\}}}_{=1}\\
&+\mu\cdot((n-1,0,[0],b)-\mathcal{L}(n,1,[0],b-1))\cdot\underbrace{1_{\{n>0\}}}_{=1}\cdot\underbrace{1_{\{\overline{k}=1\}}}_{=1}\\
&+\sum_{\widetilde{h}=1}^{H}\nu\cdot b(\widetilde{h})\cdot\Big(\mathcal{L}\Big(n,1,[\widetilde{h}],b-2\Big)-\mathcal{L}(n,1,[0],b-1)\Big)\cdot\underbrace{1_{\{k=0\}}}_{=1}\cdot\underbrace{1_{\{\overline{k}=1\}}}_{=1}\cdot\underbrace{1_{\{0<\overline{k}+k<b\}}}_{=\{b>1\}}\\
=\;&\lambda\cdot(n+1-n)+\mu\cdot\left(\left(n-1+\frac{\mu-\lambda}{2\cdot\mu}\right)-n\right)+\sum_{\widetilde{h}=1}^{H}\nu\cdot b(\widetilde{h})\cdot(n-n)\\
=\;&\lambda+\mu\cdot\left(-1+\frac{\mu-\lambda}{2\cdot\mu}\right)=\lambda-\mu+\frac{\mu-\lambda}{2}=\frac{\lambda-\mu}{2}\leq-\varepsilon.
\end{aligned}
$$

If $b > 2$, for $(n, \mathbf{k}) = \left(n, \overline{k}, [h_1, \ldots, h_k], b - (\overline{k} + k)\right)$ with $\overline{k} + k = 2, \ldots, b-1$ and $n > 0$ holds

$$
\begin{aligned}
&(\mathbf{Q} \cdot \mathcal{L})(n, \mathbf{k}) \\
=\ & \lambda \cdot (\mathcal{L}(n+1, \mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot \underbrace{1_{\{\overline{k}+k>0\}}}_{=1} \\
&+ \mu \cdot (\mathcal{L}(n-1, T_{0\searrow}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot \underbrace{1_{\{n>0\}}}_{=1} \cdot \underbrace{1_{\{\overline{k}=1\}}}_{=1} \\
&+ \sum_{\widetilde{h}=1}^{h_1-1} \nu \cdot b(\widetilde{h}) \cdot \big(\mathcal{L}(n, T_{\widetilde{h}\to 1}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})\big) \cdot \underbrace{1_{\{\widetilde{h}<h_1\}}}_{=1} \cdot \underbrace{1_{\{0<\overline{k}+k<b\}}}_{=1} \cdot \underbrace{1_{\{k>0\}}}_{=1} \\
&+ \sum_{\widetilde{h}=h_{\ell-1}}^{h_\ell-1} \nu \cdot b(\widetilde{h}) \cdot \big(\mathcal{L}(n, T_{\widetilde{h}\to \ell}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})\big) \cdot \underbrace{1_{\{h_{\ell-1}\le\widetilde{h}<h_\ell\}}}_{=1} \cdot \underbrace{1_{\{0<\overline{k}+k<b\}}}_{=1} \cdot 1_{\{k>1\}} \cdot 1_{\{\ell\in\{2,\ldots,k\}\}} \\
&+ \sum_{\widetilde{h}=h_k}^{H} \nu \cdot b(\widetilde{h}) \cdot \Big(\mathcal{L}(n, T_{\widetilde{h}\to k+1}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})\Big) \cdot \underbrace{1_{\{h_k\le\widetilde{h}\}}}_{=1} \cdot \underbrace{1_{\{0<\overline{k}+k<b\}}}_{=1} \cdot \underbrace{1_{\{k>0\}}}_{=1} \\
&+ \sum_{\ell=1}^{k} \beta \cdot (\mathcal{L}(n, T_{\ell\searrow}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot 1_{\{h_\ell=1\}} \cdot \underbrace{1_{\{k>0\}}}_{=1} \\
&+ \sum_{\ell=1}^{k} \beta \cdot (\mathcal{L}(n, T_{\ell\to 1}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot 1_{\{0<h_\ell-1<h_1\}} \cdot \underbrace{1_{\{k>0\}}}_{=1} \\
&+ \sum_{\ell=1}^{k} \beta \cdot (\mathcal{L}(n, T_{\ell\to m}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot 1_{\{0<h_{m-1}\le h_\ell-1<h_m\}} \cdot 1_{\{k>1\}} \cdot 1_{\{m\in\{2,\ldots,\ell\}\}} \\
=\ & \lambda \cdot (n+1-n) + \mu \cdot (n-1-n) \\
&+ \sum_{\widetilde{h}=1}^{h_1-1} \nu \cdot b(\widetilde{h}) \cdot (n-n) + \sum_{\widetilde{h}=h_{\ell-1}}^{h_\ell-1} \nu \cdot b(\widetilde{h}) \cdot (n-n) \cdot 1_{\{k>1\}} + \sum_{\widetilde{h}=h_k}^{H} \nu \cdot b(\widetilde{h}) \cdot (n-n) \\
&+ \sum_{\ell=1}^{k} \beta \cdot (n-n) \cdot 1_{\{h_\ell=1\}} \\
&+ \sum_{\ell=1}^{k} \beta \cdot (n-n) \cdot 1_{\{0<h_\ell-1<h_1\}} \\
&+ \sum_{\ell=1}^{k} \beta \cdot (n-n) \cdot 1_{\{0<h_{m-1}\le h_\ell-1<h_m\}} \cdot 1_{\{k>1\}} \cdot 1_{\{m\in\{2,\ldots,\ell\}\}} \\
=\ & \lambda - \mu \le \frac{\lambda-\mu}{2} \le -\varepsilon.
\end{aligned}
$$

For $(n, \mathbf{k}) = \left(n, \overline{k}, [h_1, \dots, h_k], 0\right)$ with $\overline{k} + k = b \geq 1$ and $n > 0$ holds
(Because $n > 0$ and $\overline{k} + k = b \geq 1$, we know that $\overline{k} = 1$.
Consequently, $b = 1$ implies $k = 0$, and $b > 1$ implies $k > 0$, and $b > 2$ implies $k > 1$.)

$$
\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(n, \mathbf{k}) \\
= {} & \lambda \cdot (\mathcal{L}(n+1, \mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot \underbrace{1_{\{\overline{k}+k>0\}}}_{=1} \\
& + \mu \cdot (\mathcal{L}(n-1, T_{0\searrow}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot \underbrace{1_{\{n>0\}}}_{=1} \cdot \underbrace{1_{\{\overline{k}=1\}}}_{=1} \\
& + \sum_{\ell=1}^{k} \beta \cdot (\mathcal{L}(n, T_{\ell\searrow}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot 1_{\{h_\ell=1\}} \cdot 1_{\{b>1\}} \\
& + \sum_{\ell=1}^{k} \beta \cdot (\mathcal{L}(n, T_{\ell\to 1}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot 1_{\{0<h_\ell-1<h_1\}} \cdot 1_{\{b>1\}} \\
& + \sum_{\ell=1}^{k} \beta \cdot (\mathcal{L}(n, T_{\ell\to m}\mathbf{k}) - \mathcal{L}(n, \mathbf{k})) \cdot 1_{\{0<h_{m-1}\leq h_\ell-1<h_m\}} \cdot 1_{\{m\in\{2,\dots,\ell\}\}} \cdot 1_{\{b>2\}} \\
= {} & \lambda \cdot (n+1-n) + \mu \cdot \left(\left(n - 1 + 1_{\{b=1\}} \cdot \frac{\mu-\lambda}{2\cdot\mu}\right) - n\right) \\
& + \sum_{\ell=1}^{k} \beta \cdot (n-n) \cdot 1_{\{h_\ell=1\}} \cdot 1_{\{b>1\}} \\
& + \sum_{\ell=1}^{k} \beta \cdot (n-n) \cdot 1_{\{0<h_\ell-1<h_1\}} \cdot 1_{\{b>1\}} \\
& + \sum_{\ell=1}^{k} \beta \cdot (n-n) \cdot 1_{\{0<h_{m-1}\leq h_\ell-1<h_m\}} \cdot 1_{\{m\in\{2,\dots,\ell\}\}} \cdot 1_{\{b>2\}} \\
= {} & (\lambda-\mu) \cdot 1_{\{b>1\}} + \left(\frac{\lambda-\mu}{2}\right) \cdot 1_{\{b=1\}} \leq \frac{\lambda-\mu}{2} \leq -\varepsilon.
\end{aligned}
$$

□

4.2.2.2. Properties of the stationary system

We assume throughout this section that the queueing-inventory process Z is ergodic. Some results in this section are special cases of Section D.1.2. An overview of the corresponding results is presented in Table D.1 on page 348. We provide direct proofs here to avoid the general lengthy formalism needed in queueing-environment systems.

Definition 4.2.10. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(n,\mathbf{k}) : (n,\mathbf{k}) \in E), \quad \pi(n,\mathbf{k}) := \lim_{t\to\infty} P(Z(t) = (n,\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := (\xi(n) : n \in \mathbb{N}_0), \quad \xi(n) := \lim_{t\to\infty} P(X(t) = n),$$
$$\theta := (\theta(\mathbf{k}) : \mathbf{k} \in K), \quad \theta(k) := \lim_{t\to\infty} P((Y,W)(t) = \mathbf{k}).$$

Let (X, Y, W) be a random variable which is distributed according to the queueing-inventory process in equilibrium. Therefore, X resp. (Y, W) are random variables which are distributed according to the marginal steady state probability for the production subsystem resp. for the inventory-replenishment subsystem.[2]

Proposition 4.2.11. *The queueing-inventory process Z fulfils for all $n \in \mathbb{N}_0$*

$$P(X = n, Y \neq (0,[0])) = P(X = n+1, Y \neq (0,[0])) \cdot \frac{\mu(n+1)}{\lambda(n)} \tag{4.2.3}$$

and

$$P(X = n, Y \neq (0,[0])) = P(X = 0, Y \neq (0,[0])) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \tag{4.2.4}$$

Hence, the probability that the inventory is not empty fulfils

$$P(Y \neq (0,[0])) = P(X = 0, Y \neq (0,[0])) \cdot \sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}.$$

Proof. The normalisation constant, as the sum of the solution $\mathbf{x} = (x(z) : z \in E)$ of the global balance equation $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ for Z, has to be finite because the queueing-inventory process Z is positive recurrent. Then with

$$\frac{x(n,\mathbf{k})}{\sum_{n=0}^{\infty}\sum_{\mathbf{k}\in K} x(n,\mathbf{k})} = \pi(n,\mathbf{k}) = P\left(X = n, Y = \left(\overline{k}, [h_1,\ldots,h_k]\right)\right)$$

follows in steady state from (4.2.1)

$$\underbrace{\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} \pi(n,\mathbf{k})}_{=P(X=n,Y\neq(0,[0]))} = \underbrace{\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} \pi(n+1,\mathbf{k})}_{=P(X=n+1,Y\neq(0,[0]))} \cdot \frac{\mu(n+1)}{\lambda(n)}$$

[2] It should be noted that $\pi(n,\mathbf{k}) = \pi(n,\overline{k},[h_1,\ldots,h_k], b-k) = P(X = n, Y = \left(\overline{k},[h_1,\ldots,h_k]\right), W = b-k) = P(X = n, Y = \left(\overline{k},[h_1,\ldots,h_k]\right))$, because the base stock level b is a fixed parameter.

and from (4.2.2)

$$\underbrace{\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} \pi(n,\mathbf{k})}_{=P(X=n,Y\neq(0,[0]))} = \underbrace{\sum_{\mathbf{k}\in K\setminus\{(0,[0],b)\}} \pi(0,\mathbf{k})}_{=P(X=0,Y\neq(0,[0]))} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \tag{4.2.5}$$

Summation of $P(X=n, Y\neq(0,[0]))$ over $n\in\mathbb{N}_0$ yields

$$\begin{aligned} P(Y\neq(0,[0])) &= \sum_{n=0}^{\infty} P(X=n, Y\neq(0,[0])) \\ &\overset{(4.2.4)}{=} \sum_{n=0}^{\infty} P(X=0, Y\neq(0,[0])) \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \\ &= P(X=0, Y\neq(0,[0]))\cdot \sum_{n=0}^{\infty}\prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \end{aligned}$$

□

Corollary 4.2.12. *For the conditional distribution of the queue length process conditioned on $\{Y\neq(0,[0])\}$ holds for $n\in\mathbb{N}_0$*

$$P\left(X=n|Y\neq(0,[0])\right) = P\left(X=0|Y\neq(0,[0])\right)\cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}$$

with

$$P\left(X=0|Y\neq(0,[0])\right) = \left(\sum_{n=0}^{\infty}\prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}\right)^{-1}.$$

This shows that the conditional queue length process under the condition that the inventory is not empty has in equilibrium the same structure as a birth-and-death process with queue-length-dependent intensities.

Proof. $P\left(Y\neq(0,[0]\right)>0$ because of ergodicity and equation (4.2.4) imply for $n\in\mathbb{N}_0$

$$\begin{aligned} P\left(X=n|Y\neq(0,[0])\right) &= \frac{P(X=n, Y\neq(0,[0]))}{P\left(Y\neq(0,[0])\right)} \\ &\overset{(4.2.4)}{=} \frac{P(X=0, Y\neq(0,[0]))}{P\left(Y\neq(0,[0])\right)}\cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \\ &= P\left(X=0|Y\neq(0,[0])\right)\cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \end{aligned}$$

Hence,

$$P\left(X=0|Y\neq(0,[0])\right) = \left(\sum_{n=0}^{\infty}\prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}\right)^{-1}.$$

□

Remark 4.2.13. The limiting and stationary distribution of Z is in general not of product form (4.2.6). From Corollary 4.2.12 follows that if the stationary distribution has a product form

$$\pi(n, \mathbf{k}) = \xi(n) \cdot \theta(\mathbf{k}), \qquad n \in \mathbb{N}_0,\ \mathbf{k} \in K, \tag{4.2.6}$$

then

$$\xi(n) = C^{-1} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}, \qquad n \in \mathbb{N}_0,$$

with normalisation constant

$$C = \sum_{\tilde{n}=0}^{\infty} \prod_{m=1}^{\tilde{n}} \frac{\lambda(m-1)}{\mu(m)}.$$

By substitution of this stationary distribution into the global balance equations we see that the limiting and stationary distribution of Z in general cannot be of product form. In Proposition 4.2.29 we show this by contradiction for the production-inventory system with exponentially distributed life times which is a special case of the system with phase-type distributed life times.

Special case: Queue-length-independent arrival and service rates

In this paragraph, we analyse the queueing-inventory system with state-independent service rates μ and arrival rates λ. Recall that the queueing-inventory process Z is ergodic.

The following proposition is a special case of Proposition 4.2.11.

Proposition 4.2.14. *If the queueing-inventory process Z is positive recurrent, in steady state it fulfils for all $n \in \mathbb{N}_0$*

$$P(X = n, Y \neq (0, [0])) = P(X = n+1, Y \neq (0, [0])) \cdot \frac{\mu}{\lambda} \tag{4.2.7}$$

and

$$P(X = n, Y \neq (0, [0])) = P(X = 0, Y \neq (0, [0])) \cdot \left(\frac{\lambda}{\mu}\right)^n. \tag{4.2.8}$$

The following proposition is a special case of Corollary 4.2.12.

Corollary 4.2.15. *For the conditional distribution of the queue length process conditioned on $\{Y \neq (0, [0])\}$ holds for $n \in \mathbb{N}_0$*

$$P(X = n | Y \neq (0, [0])) = P(X = 0 | Y \neq (0, [0])) \cdot \left(\frac{\lambda}{\mu}\right)^n$$

with

$$P(X = 0 | Y \neq (0, [0])) = \left(1 - \frac{\lambda}{\mu}\right).$$

This shows that the conditional queue length process under the condition that the inventory is not empty has in equilibrium the same structure as a birth-and-death process with birth-rates λ and death-rates μ.

Proposition 4.2.16. *For the queueing-inventory process Z holds the following equilibrium of probability flows*

$$\underbrace{P(Y \neq (0,[0])) \cdot \lambda}_{\text{effective arrival rate}} = \underbrace{P(X > 0, Y \neq (0,[0])) \cdot \mu}_{\text{effective departure rate}}.$$

Hence, the probability that the inventory is not empty fulfils

$$P(Y \neq (0,[0])) = P(X > 0, Y \neq (0,[0])) \cdot \frac{\mu}{\lambda}$$

and

$$P(Y \neq (0,[0])) = P(X = 0, Y \neq (0,[0])) \cdot \frac{\mu}{\mu - \lambda}.$$

Remark 4.2.17. The effective departure rate is usually called throughput.

The loss rate is given by $\lambda \cdot P\left(Y = (0,[0])\right)$.

Proof. Summation of $P(X = n, Y \neq (0,[0]))$ over $n \in \mathbb{N}_0$ yields

$$\begin{aligned} P(Y \neq (0,[0])) &= \sum_{n=0}^{\infty} P(X = n, Y \neq (0,[0])) \overset{(4.2.7)}{=} \sum_{n=0}^{\infty} P(X = n+1, Y \neq (0,[0])) \cdot \frac{\mu}{\lambda} \\ &= P(X > 0, Y \neq (0,[0])) \cdot \frac{\mu}{\lambda}. \end{aligned}$$

and

$$\begin{aligned} P(Y \neq (0,[0])) &= \sum_{n=0}^{\infty} P(X = n, Y \neq (0,[0])) \overset{(4.2.8)}{=} \sum_{n=0}^{\infty} P(X = 0, Y \neq (0,[0])) \cdot \left(\frac{\lambda}{\mu}\right)^n \\ &= P(X = 0, Y \neq (0,[0])) \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n = P(X = 0, Y \neq (0,[0])) \cdot \frac{\mu}{\mu - \lambda}. \end{aligned}$$

□

Remark 4.2.18. The limiting and stationary distribution of Z is in general not of product form. This can be proven similarly to the proof of Proposition 4.2.35 for the production-inventory system with queue-length independent arrival and service rates. From Corollary 4.2.15 follows that if the stationary distribution has a product form

$$\pi\left(n, \mathbf{k}\right) = \xi(n) \cdot \theta\left(\mathbf{k}\right), \qquad n \in \mathbb{N}_0,\ \mathbf{k} \in K,$$

then

$$\xi(n) = C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n, \qquad n \in \mathbb{N}_0,$$

with normalisation constant $C^{-1} = \left(1 - \frac{\lambda}{\mu}\right)$.

By substitution of this stationary distribution into the global balance equations we see that the limiting and stationary distribution of Z in general cannot be of product form. In Proposition 4.2.29 we have shown this by contradiction for the production-inventory system with exponentially distributed life times which is a special case of the system with phase-type distributed life times.

4.2.3. Exponentially distributed life time

As mentioned before by varying the parameters of the phase-type distribution any distribution on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$ can be approximated sufficiently close. A phase-type distribution with one phase is an exponential distribution. In this section, we present the results for the supply chain with a single location ($J = 1$, $\overline{J} = \{1\}$) as described in Section 4.2.1, where the life time of each raw material in the inventory is exponentially distributed with rate $\gamma > 0$. We call γ the ageing rate. The item of raw material that is in the production process cannot perish. Consequently, we will take into consideration whether there are customers in the system:

- If there is at least one customer at the location and there are $k > 0$ items of raw material in the inventory, the production server is working. Therefore, an item of raw material is in production, which cannot perish, and so the loss rate of inventory due to perishing is $\gamma \cdot (k-1)$.

- If there are no customers at the location and $k > 0$ items of raw material in the inventory, then the loss rate of inventory due to perishing is $\gamma \cdot k$.

We call the functions $k \mapsto \gamma \cdot k$ resp. $k \mapsto \gamma \cdot (k-1)_+$ with $(k-1)_+ := \max(0, k-1)$ ageing regimes. The different ageing regimes determine the state-dependent overall loss rates of inventory due to perishing.

An overview of the corresponding results is presented in Table D.1 on page 348.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X(t)$ the number of customers present at time $t \geq 0$ either waiting or in service (queue length). By $Y(t)$ we denote the size of the inventory at time $t \geq 0$. By $W(t)$ we denote the number of replenishment orders at the supplier at time $t \geq 0$ either waiting or in service (queue length).

We define the joint queueing-inventory process of this system by

$$Z = ((X(t), Y(t), W(t)) \,:\, t \geq 0)\,.$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process. The state space of Z is

$$E = \{(n, k, b-k) : n \in \mathbb{N}_0,\, k \in \{0, \dots, b\}\}\,.$$

Note the redundancy in the state space: $W(t) = b - Y(t)$. We prefer to carry all information explicitly with because the dynamics of the system are easier visible.

4.2.3.1. Ergodicity

The stochastic queueing-inventory process Z is a homogeneous strong Markov process and has an infinitesimal generator $\mathbf{Q} = (q(z; \tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(n,k,b-k)\in E$:

$$\begin{aligned}
q\left((n,k,b-k);(n+1,k,b-k)\right) &= \lambda(n)\cdot 1_{\{k>0\}},\\
q\left((n,k,b-k);(n,k-1,b-k+1)\right) &= \left(\gamma\cdot(k-1)\cdot 1_{\{n>0\}}+\gamma\cdot k\cdot 1_{\{n=0\}}\right)1_{\{k>0\}},\\
q\left((n,k,b-k);(n-1,k-1,b-k+1)\right) &= \mu(n)\cdot 1_{\{n>0\}}\cdot 1_{\{k>0\}},\\
q\left((n,k,b-k);(n,k+1,b-k-1)\right) &= \nu\cdot 1_{\{k<b\}}.
\end{aligned}$$

Furthermore, $q(z;\tilde z)=0$ for any other pair $z\neq\tilde z$, and

$$q\left(z;z\right)=-\sum_{\substack{\tilde z\in E,\\ z\neq\tilde z}} q\left(z;\tilde z\right)\qquad\forall z\in E.$$

From the above transition rates follows similarly as in Appendix E.1 that Z is irreducible on the state space E.

The following proposition is a special case of Proposition 4.2.4.

Proposition 4.2.19. *If the queueing-inventory process Z is recurrent, then any solution $\mathbf{x}=(x\left(z\right):z\in E)$ of the global balance equations $\mathbf{x}\cdot\mathbf{Q}=\mathbf{0}$ fulfils for all $n\in\mathbb{N}_0$*

$$\sum_{k=1}^{b} x(n,k,b-k)=\sum_{k=1}^{b} x(n+1,k,b-k)\cdot\frac{\mu(n+1)}{\lambda(n)}$$

and

$$\sum_{k=1}^{b} x(n,k,b-k)=\sum_{k=1}^{b} x(0,k,b-k)\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}.$$

The following proposition is a special case of Proposition 4.2.5.

Proposition 4.2.20. *If the queueing-inventory process Z is ergodic, it holds $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}<\infty$.*

The following proposition is a special case of Proposition 4.2.6.

Proposition 4.2.21.

(a) *The queueing-inventory process Z is ergodic if for an $M/M/1/\infty$ queue with queue-length dependent arrival intensities $\lambda(n)>0$ and service intensities $\mu(n)>0$ there exists a Lyapunov function $\widetilde{\mathcal{L}}:\mathbb{N}_0\to\mathbb{R}_0^+$ with finite exception set $\widetilde F$ and constant $\widetilde\varepsilon>0$, which satisfies the Foster-Lyapunov stability criterion, and $\sup_{n\in\mathbb{N}_0}\mu(n)<\infty$.*

(b) *The queueing-inventory process Z is ergodic if there exists $N\in\mathbb{N}_0$ such that $\inf_{n\geq N}\left(\mu(n)-\lambda(n)\right)>0$ and $\sup_{n\in\mathbb{N}_0}\mu(n)<\infty$.*

Remark 4.2.22. The condition $\sup_{n\in\mathbb{N}_0}\mu(n)<\infty$ can be weakened by $\inf_{n\in\mathbb{N}_0}\widehat{c}_n>0$, where $\widehat{c}_n$ is defined in Lemma D.1.7.

Remark 4.2.23. The queueing-inventory system can be modelled as a level-dependent quasi-birth-and-death process (LDQBD process). Under the assumptions from the above proposition, the queueing-inventory system is ergodic and hence, we can use the algorithm of Bright and Taylor [BT95] to calculate the equilibrium distributions in LDQBD processes.

Special case: Queue-length-independent arrival and service rates
In this paragraph, we analyse the queueing-inventory system with state-independent service rates μ and arrival rates λ in view of a sufficient and necessary criterion for ergodicity.

The following theorem is a special case of Theorem 4.2.9.

Theorem 4.2.24. *The queueing-inventory process Z is ergodic if and only if $\lambda < \mu$.*

4.2.3.2. Properties of the stationary system

In this section, we assume that the queueing-inventory process Z is ergodic.

Definition 4.2.25. For the queueing-inventory process Z in a state space E, whose limiting distribution exists, we define

$$\pi := (\pi(n,k,b-k) : (n,k,b-k) \in E), \quad \pi(n,k,b-k) := \lim_{t\to\infty} P(Z(t) = (n,k,b-k))$$

and the appropriate marginal distributions

$$\xi := (\xi(n) : n \in \mathbb{N}_0), \quad \xi(n) := \lim_{t\to\infty} P(X(t) = n),$$

$$\theta := (\theta(k) : k \in K), \quad \theta(k) := \lim_{t\to\infty} P((Y,W)(t) = (k,b-k)).$$

Let (X,Y,W) be a random variable which is distributed according to the queueing-inventory process in equilibrium. Therefore, X resp. (Y,W) are random variables which are distributed according to the marginal steady state probability for the production subsystem resp. for the replenishment-inventory subsystem.[3]

The following proposition is a special case of Proposition 4.2.11.

Proposition 4.2.26. *The queueing-inventory process Z fulfils for all $n \in \mathbb{N}_0$*

$$P(X=n, Y>0) = P(X=n+1, Y>0) \cdot \frac{\mu(n+1)}{\lambda(n)}$$

and

$$P(X=n, Y>0) = P(X=0, Y>0) \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}.$$

Hence, the probability that the inventory is not empty fulfils

$$P(Y>0) = P(X=0, Y>0) \cdot \sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}.$$

[3]It should be noted that $\pi(n,k,b-k) = P(X=n, Y=k, W=b-k) = P(X=n, Y=k)$ because the base stock level b is a fixed parameter.

The following corollary is a special case of Corollary 4.2.12.

Corollary 4.2.27. *For the conditional distribution of the queue length process conditioned on $\{Y > 0\}$ holds for $n \in \mathbb{N}_0$*

$$P(X = n|Y > 0) = P(X = 0|Y > 0) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}$$

with

$$P(X = 0|Y > 0) = \left(\sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \right)^{-1}.$$

This shows that the conditional queue length process under the condition that the inventory is not empty has in equilibrium the same structure as a birth-and-death process with queue-length-dependent intensities.

Proposition 4.2.28. *For the inventory process holds for $\ell = 1, \ldots, b$*

$$\begin{aligned} P(Y = \ell - 1) \cdot \nu = P(X = 0, Y = \ell) \cdot \gamma \cdot \ell + P(X > 0, Y = \ell) \cdot \gamma \cdot (\ell - 1) \\ + \sum_{n=1}^{\infty} P(X = n, Y = \ell) \cdot \mu(n). \end{aligned} \tag{4.2.9}$$

Hence, the probability that a replenishment order is outstanding fulfils

$$P(Y < b) = \frac{1}{\nu} \cdot \left[\gamma \cdot E(Y) - P(X > 0, Y > 0) \cdot \gamma + \sum_{\ell=1}^{b} \sum_{n=1}^{\infty} P(X = n, Y = \ell) \cdot \mu(n) \right].$$

Proof. The equation can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a). For $\ell \in \{1, \ldots, b\}$, the equation can be proven by a cut, which divides E into complementary sets according to the size of the inventory that is less than or equal to $\ell - 1$ or greater than $\ell - 1$, i.e. into the sets

$$\Big\{(n, k, b-k) : n \in \mathbb{N}_0,\ k \in \{0, \ldots, \ell - 1\}\Big\},$$

$$\Big\{(\widetilde{n}, \widetilde{k}, b - \widetilde{k}) : \widetilde{n} \in \mathbb{N}_0,\ \widetilde{k} \in \{\ell, \ldots, b\}\Big\}, \quad \ell \in \{1, \ldots, b\}.$$

Then, it follows for $\ell \in \{1, \ldots, b\}$

$$\begin{aligned} &\sum_{n=0}^{\infty} \sum_{k=0}^{\ell-1} \sum_{\widetilde{n}=0}^{\infty} \sum_{\widetilde{k}=\ell}^{b} \pi(n, k, b-k) \cdot q((n, k, b-k); (\widetilde{n}, \widetilde{k}, b - \widetilde{k})) \\ &= \sum_{\widetilde{n}=0}^{\infty} \sum_{\widetilde{k}=\ell}^{b} \sum_{n=0}^{\infty} \sum_{k=0}^{\ell-1} \pi(\widetilde{n}, \widetilde{k}, b - \widetilde{k}) \cdot q((\widetilde{n}, \widetilde{k}, b - \widetilde{k}); (n, k, b-k)) \\ \Leftrightarrow \quad &\sum_{n=0}^{\infty} \pi(n, \ell - 1, b - \ell + 1) \cdot \nu \\ &= \pi(0, \ell, b - \ell) \cdot \gamma \cdot \ell + \sum_{n=1}^{\infty} \pi(n, \ell, b-\ell) \cdot \gamma \cdot (\ell - 1) + \sum_{n=1}^{\infty} \pi(n, \ell, b - \ell) \cdot \mu(n) \end{aligned}$$

$$
\begin{aligned}
\Leftrightarrow \quad & \underbrace{\sum_{n=0}^{\infty} \pi(n, \ell-1, b-\ell+1)\cdot\nu}_{=P(Y=\ell-1)} \\
&= \underbrace{\pi(0,\ell,b-\ell)}_{=P(X=0,Y=\ell)}\cdot\gamma\cdot\ell + \underbrace{\sum_{n=1}^{\infty}\pi(n,\ell,b-\ell)}_{=P(X>0,Y=\ell)}\cdot\gamma\cdot(\ell-1) + \sum_{n=1}^{\infty}\underbrace{\pi(n,\ell,b-\ell)}_{=P(X=n,Y=\ell)}\cdot\mu(n).
\end{aligned}
$$

Then, it follows from (4.2.9)

$$
\begin{aligned}
& P(Y<b)\cdot\nu = \sum_{\ell=0}^{b-1} P(Y=\ell)\cdot\nu = \sum_{\ell=1}^{b} P(Y=\ell-1)\cdot\nu \\
&= \sum_{\ell=1}^{b} P(X=0,Y=\ell)\cdot\gamma\cdot\ell + \sum_{\ell=1}^{b} P(X>0,Y=\ell)\cdot\gamma\cdot(\ell-1) \\
&\quad + \sum_{\ell=1}^{b}\sum_{n=1}^{\infty} P(X=n,Y=\ell)\cdot\mu(n) \\
&= \sum_{\ell=1}^{b} P(X=0,Y=\ell)\cdot\gamma\cdot\ell + \sum_{\ell=1}^{b} P(X>0,Y=\ell)\cdot\gamma\cdot\ell - \sum_{\ell=1}^{b} P(X>0,Y=\ell)\cdot\gamma \\
&\quad + \sum_{\ell=1}^{b}\sum_{n=1}^{\infty} P(X=n,Y=\ell)\cdot\mu(n) \\
&= \sum_{\ell=1}^{b} P(Y=\ell)\cdot\gamma\cdot\ell - \sum_{\ell=1}^{b} P(X>0,Y=\ell)\cdot\gamma + \sum_{\ell=1}^{b}\sum_{n=1}^{\infty} P(X=n,Y=\ell)\cdot\mu(n) \\
&= \gamma\cdot\underbrace{\sum_{\ell=0}^{b} P(Y=\ell)\cdot\ell}_{=E(Y)} - P(X>0,Y>0)\cdot\gamma + \sum_{\ell=1}^{b}\sum_{n=1}^{\infty} P(X=n,Y=\ell)\cdot\mu(n) \\
&= \gamma\cdot E(Y) - P(X>0,Y>0)\cdot\gamma + \sum_{\ell=1}^{b}\sum_{n=1}^{\infty} P(X=n,Y=\ell)\cdot\mu(n).
\end{aligned}
$$

□

Proposition 4.2.29. *The limiting and stationary distribution of the queueing-inventory process Z is in general not of product form.*

Proof. If the stationary distribution has a product form, it holds for any $n \in \mathbb{N}_0$

$$P\left(X=n, Y>0\right)=P\left(X=n\right)\cdot P\left(Y>0\right).$$

Then, it follows from Corollary 4.2.27

$$\begin{aligned}P(X=n) &= P\left(X=0|Y>0\right)\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}\\ &= \left(\sum_{\widetilde{n}=0}^{\infty}\prod_{m=1}^{\widetilde{n}}\frac{\lambda(m-1)}{\mu(m)}\right)^{-1}\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}, \qquad n\in\mathbb{N}_0.\end{aligned}$$

Consequently, if the stationary distribution has a product form

$$\pi\left(n,k,b-k\right)=\xi(n)\cdot\theta\left(k,b-k\right), \qquad n\in\mathbb{N}_0,\ k\in\{0,1,\ldots,b\},$$

then

$$\xi(n)=C^{-1}\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}, \qquad n\in\mathbb{N}_0,$$

with normalisation constant

$$C=\sum_{\widetilde{n}=0}^{\infty}\prod_{m=1}^{\widetilde{n}}\frac{\lambda(m-1)}{\mu(m)}.$$

It has to be shown that this distribution does not satisfy the global balance equations

$$\begin{aligned}&\pi\left(n,k,b-k\right)\cdot\Big(\Big(\lambda(n)+\gamma\cdot(k-1)\cdot 1_{\{n>0\}}+\gamma\cdot k\cdot 1_{\{n=0\}}\Big)\cdot 1_{\{k>0\}}\\ &\qquad\qquad +\mu(n)\cdot 1_{\{n>0\}}\cdot 1_{\{k>0\}}+\nu\cdot 1_{\{k<b\}}\Big)\\ &=\pi\left(n-1,k,b-k\right)\cdot\lambda(n-1)\cdot 1_{\{n>0\}}\cdot 1_{\{k>0\}}\\ &\quad+\pi\left(n+1,k+1,b-k-1\right)\cdot\mu(n+1)\cdot 1_{\{k<b\}}\\ &\quad+\pi\left(n,k+1,b-k-1\right)\cdot\Big(\gamma\cdot k\cdot 1_{\{n>0\}}+\gamma\cdot(k+1)\cdot 1_{\{n=0\}}\Big)\cdot 1_{\{k<b\}}\\ &\quad+\pi\left(n,k-1,b-k+1\right)\cdot\nu\cdot 1_{\{k>0\}}.\end{aligned}$$

Substitution of $\pi\left(n,k,b-k\right)=\xi(n)\cdot\theta\left(k,b-k\right)$ into the global balance equations directly leads to

$$\begin{aligned}&\xi(n)\cdot\theta\left(k,b-k\right)\cdot\Big(\Big(\lambda(n)+\gamma\cdot(k-1)\cdot 1_{\{n>0\}}+\gamma\cdot k\cdot 1_{\{n=0\}}\Big)\cdot 1_{\{k>0\}}\\ &\qquad\qquad +\mu(n)\cdot 1_{\{n>0\}}\cdot 1_{\{k>0\}}+\nu\cdot 1_{\{k<b\}}\Big)\\ &=\xi(n-1)\cdot\theta\left(k,b-k\right)\cdot\lambda(n-1)\cdot 1_{\{n>0\}}\cdot 1_{\{k>0\}}\\ &\quad+\xi(n+1)\cdot\theta\left(k+1,b-k-1\right)\cdot\mu(n+1)\cdot 1_{\{k<b\}}\\ &\quad+\xi(n)\cdot\theta\left(k+1,b-k-1\right)\cdot\Big(\gamma\cdot k\cdot 1_{\{n>0\}}+\gamma\cdot(k+1)\cdot 1_{\{n=0\}}\Big)\cdot 1_{\{k<b\}}\\ &\quad+\xi(n)\cdot\theta\left(k-1,b-k+1\right)\cdot\nu\cdot 1_{\{k>0\}}.\end{aligned}$$

By substitution of $\xi(n) = C^{-1} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}$ we obtain

$$\begin{aligned}
&\xi(n) \cdot \theta\left(k, b-k\right) \cdot \Big(\Big(\lambda(n) + \gamma \cdot (k-1) \cdot 1_{\{n>0\}} + \gamma \cdot k \cdot 1_{\{n=0\}} \Big) \cdot 1_{\{k>0\}} \\
&\qquad + \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k<b\}} \Big) \\
&= \xi(n) \cdot \theta\left(k, b-k\right) \cdot \frac{\mu(n)}{\lambda(n-1)} \cdot \lambda(n-1) \cdot 1_{\{n>0\}} \cdot 1_{\{k>0\}} \\
&+ \xi(n) \cdot \theta\left(k+1, b-k-1\right) \cdot \frac{\lambda(n)}{\mu(n+1)} \cdot \mu(n+1) \cdot 1_{\{k<b\}} \\
&+ \xi(n) \cdot \theta\left(k+1, b-k-1\right) \cdot \Big(\gamma \cdot k \cdot 1_{\{n>0\}} + \gamma \cdot (k+1) \cdot 1_{\{n=0\}} \Big) \cdot 1_{\{k<b\}} \\
&+ \xi(n) \cdot \theta\left(k-1, b-k+1\right) \cdot \nu \cdot 1_{\{k>0\}}.
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
&\xi(n) \cdot \theta\left(k, b-k\right) \cdot \Big(\Big(\lambda(n) + \gamma \cdot (k-1) \cdot 1_{\{n>0\}} + \gamma \cdot k \cdot 1_{\{n=0\}} \Big) \cdot 1_{\{k>0\}} \\
&\qquad + \mu \cdot 1_{\{n>0\}} \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k<b\}} \Big) \\
&= \xi(n) \cdot \theta\left(k, b-k\right) \cdot \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k>0\}} \\
&+ \xi(n) \cdot \theta\left(k+1, b-k-1\right) \cdot \lambda(n) \cdot 1_{\{k<b\}} \\
&+ \xi(n) \cdot \theta\left(k+1, b-k-1\right) \cdot \Big(\gamma \cdot k \cdot 1_{\{n>0\}} + \gamma \cdot (k+1) \cdot 1_{\{n=0\}} \Big) \cdot 1_{\{k<b\}} \\
&+ \xi(n) \cdot \theta\left(k-1, b-k+1\right) \cdot \nu \cdot 1_{\{k>0\}}.
\end{aligned}$$

Cancelling $\xi(n)$ and the sum with the terms $\mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k>0\}}$ on both sides of the equation leads to

$$\begin{aligned}
&\theta\left(k, b-k\right) \cdot \Big(\Big(\lambda(n) + \gamma \cdot (k-1) \cdot 1_{\{n>0\}} + \gamma \cdot k \cdot 1_{\{n=0\}} \Big) \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k<b\}} \Big) \\
&= \theta\left(k+1, b-k-1\right) \cdot \lambda(n) \cdot 1_{\{k<b\}} \\
&+ \theta\left(k+1, b-k-1\right) \cdot \Big(\gamma \cdot k \cdot 1_{\{n>0\}} + \gamma \cdot (k+1) \cdot 1_{\{n=0\}} \Big) \cdot 1_{\{k<b\}} \\
&+ \theta\left(k-1, b-k+1\right) \cdot \nu \cdot 1_{\{k>0\}}.
\end{aligned} \tag{4.2.10}$$

However, this stands in contradiction to the product form assumption since in general $\theta(k, b-k)$ cannot be defined independently of n. □

Special case: Queue-length-independent arrival and service rates
In this paragraph, we analyse the queueing-inventory system with state-independent service rates μ and arrival rates λ. Recall that the queueing-inventory process Z is ergodic.

The following proposition is a special case of Proposition 4.2.26.

Proposition 4.2.30. *The queueing-inventory process Z fulfils for all $n \in \mathbb{N}_0$*

$$P(X = n, Y > 0) = P(X = n + 1, Y > 0) \cdot \frac{\mu}{\lambda} \tag{4.2.11}$$

and

$$P(X = n, Y > 0) = P(X = 0, Y > 0) \cdot \left(\frac{\lambda}{\mu}\right)^n. \tag{4.2.12}$$

The following corollary is a special case of Corollary 4.2.27.

Corollary 4.2.31. *For the conditional distribution of the queue length process conditioned on $\{Y > 0\}$ holds for $n \in \mathbb{N}_0$*

$$P\left(X = n | Y > 0\right) = P\left(X = 0 | Y > 0\right) \cdot \left(\frac{\lambda}{\mu}\right)^n$$

with

$$P\left(X = 0 | Y > 0\right) = \left(1 - \frac{\lambda}{\mu}\right).$$

This shows that the conditional queue length process under the condition that the inventory is not empty has in equilibrium the same structure as a birth-and-death process with birth-rates λ and death-rates μ.

Proposition 4.2.32. *For the queueing-inventory process Z holds the following equilibrium of probability flows*

$$\underbrace{P(Y > 0) \cdot \lambda}_{\text{effective arrival rate}} = \underbrace{P(X > 0, Y > 0) \cdot \mu}_{\text{effective departure rate}}.$$

Hence, the probability that the inventory is not empty fulfils

$$P(Y > 0) = P(X > 0, Y > 0) \cdot \frac{\mu}{\lambda} \tag{4.2.13}$$

and

$$P(Y > 0) = P(X = 0, Y > 0) \cdot \frac{\mu}{\mu - \lambda}. \tag{4.2.14}$$

Remark 4.2.33. The effective departure rate is usually called throughput.
The loss rate is given by $\lambda \cdot P\left(Y = 0\right)$.

Proof. Summation of $P(X = n, Y > 0)$ over $n \in \mathbb{N}_0$ yields

$$\begin{aligned} P(Y > 0) = \sum_{n=0}^{\infty} P(X = n, Y > 0) &\overset{(4.2.11)}{=} \sum_{n=0}^{\infty} P(X = n + 1, Y > 0) \cdot \frac{\mu}{\lambda} \\ &= P(X > 0, Y > 0) \cdot \frac{\mu}{\lambda} \end{aligned}$$

and

$$\begin{aligned}P(Y>0) &= \sum_{n=0}^{\infty} P(X=n, Y>0) \overset{(4.2.12)}{=} \sum_{n=0}^{\infty} P(X=0, Y>0) \cdot \left(\frac{\lambda}{\mu}\right)^n \\ &= P(X=0, Y>0) \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n = P(X=0, Y>0) \cdot \frac{\mu}{\mu-\lambda}.\end{aligned}$$

□

The following proposition is a special case of Proposition 4.2.28.

Proposition 4.2.34. *For the inventory process holds for $\ell = 1, \dots, b$*

$$\begin{aligned}P(Y=\ell-1) \cdot \nu &= P(X=0, Y=\ell) \cdot \gamma \cdot \ell + P(X>0, Y=\ell) \cdot \big(\mu + \gamma \cdot (\ell-1)\big) \\ &= P(X=0, Y=\ell) \cdot (\gamma-\mu) + P(Y=\ell) \cdot \big(\mu + \gamma \cdot (\ell-1)\big).\end{aligned}$$

The probability that a replenishment order is outstanding fulfils

$$P(Y<b) = \frac{1}{\nu} \cdot \left[E\left(Y\right) \cdot \gamma + P(X>0, Y>0) \cdot (\mu-\gamma)\right].$$

Proposition 4.2.35. *The limiting and stationary distribution of the queueing-inventory process Z is in general not of product form.*

Proof. The structure of the proof is similar to the proof of Proposition 4.2.29 for the production-inventory system with queue-length-dependent arrival and service rates. If the stationary distribution has a product form, it holds for any $n \in \mathbb{N}_0$

$$P\left(X=n, Y>0\right) = P\left(X=n\right) \cdot P\left(Y>0\right).$$

Then, it follows from Corollary 4.2.31 for $n \in \mathbb{N}_0$

$$\begin{aligned}P(X=n) &= \frac{P\left(X=n\right) \cdot P\left(Y>0\right)}{P\left(Y>0\right)} = \frac{P\left(X=n, Y>0\right)}{P\left(Y>0\right)} = P\left(X=n | Y>0\right) \\ &= P\left(X=0 | Y>0\right) \cdot \left(\frac{\lambda}{\mu}\right)^n = \left(1-\frac{\lambda}{\mu}\right) \cdot \left(\frac{\lambda}{\mu}\right)^n.\end{aligned}$$

Consequently, if the stationary distribution has a product form

$$\pi\left(n, k, b-k\right) = \xi(n) \cdot \theta\left(k, b-k\right), \qquad n \in \mathbb{N}_0,\ k \in \{0, 1, \dots, b\},$$

then

$$\xi(n) = C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n, \qquad n \in \mathbb{N}_0,$$

with normalisation constant $C^{-1} = \left(1 - \frac{\lambda}{\mu}\right)$.

By substitution of this stationary distribution into the global balance equations we also get an equation as in (4.2.10) (with λ instead of $\lambda(n)$) which is in contradiction to the product form assumption. □

Queue-length-independent arrival and service rates and base stock level $b = 1$
In this paragraph, we analyse the queueing-inventory system with state-independent service rates μ and arrival rates λ and base stock level $b = 1$. Recall that the queueing-inventory process Z is ergodic.

Proposition 4.2.36. *Consider a queueing-inventory system with base stock level $b = 1$ and queue-length-independent service rates μ and arrival rates λ. The limiting and stationary distribution of the queueing-inventory process $Z = ((X(t), Y(t), W(t)) : t \geq 0)$ is*

$$\begin{aligned}
\pi(0,0,1) &= C^{-1} \cdot \frac{\lambda+\gamma}{\nu}, \\
\pi(n,0,1) &= C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n \cdot \frac{\lambda}{\nu}, && n > 0, \\
\pi(n,1,0) &= C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n, && n \geq 0, \qquad (4.2.15)
\end{aligned}$$

with normalisation constant

$$C = \frac{\mu}{\mu-\lambda} \cdot \left(1 + \frac{\lambda}{\nu}\right) + \frac{\gamma}{\nu}. \qquad (4.2.16)$$

Proof. The stochastic queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(n, k, b-k) \in E$:

$$\begin{aligned}
q\left((n,1,0);(n+1,1,0)\right) &= \lambda, \\
q\left((n,1,0);(n,0,1)\right) &= \gamma \cdot 1_{\{n=0\}}, \\
q\left((n,1,0);(n-1,0,1\right) &= \mu \cdot 1_{\{n>0\}}, \\
q\left((n,0,1);(n,1,0)\right) &= \nu.
\end{aligned}$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

Therefore, the global balance equations $\pi \cdot \mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are for $(n, k, b-k) \in E$ given by

$$\begin{aligned}
&\pi\left(n,k,b-k\right) \cdot \left(\left(\lambda + \gamma \cdot 1_{\{n=0\}}\right) \cdot 1_{\{k=1\}} + \mu \cdot 1_{\{n>0\}} \cdot 1_{\{k=1\}} + \nu \cdot 1_{\{k=0\}}\right) \\
= &\pi\left(n-1,k,b-k\right) \cdot \lambda \cdot 1_{\{n>0\}} \cdot 1_{\{k=1\}} \\
&+\pi\left(n+1,k+1,b-1\right) \cdot \mu \cdot 1_{\{k=0\}} \\
&+\pi\left(n,k+1,b-k-1\right) \cdot \gamma \cdot 1_{\{n=0\}} \cdot 1_{\{k=0\}} \\
&+\pi\left(n,k-1,b-k+1\right) \cdot \nu \cdot 1_{\{k=1\}}.
\end{aligned}$$

It has to be shown that the distribution satisfies these global balance equations.
For $n = 0$ and $k = 0$ holds

$$\begin{aligned}
&\pi\left(0,0,1\right) \cdot \nu = \pi\left(1,1,0\right) \cdot \mu + \pi\left(0,1,0\right) \cdot \gamma \\
\Leftrightarrow\quad &\frac{\lambda+\gamma}{\nu} \cdot \nu = \left(\frac{\lambda}{\mu}\right) \cdot \mu + \gamma \quad\Leftrightarrow\quad \lambda + \gamma = \lambda + \gamma.
\end{aligned}$$

For $n = 0$ and $k = 1$ holds

$$\begin{aligned} &\pi(0,1,0) \cdot \Big(\lambda + \gamma\Big) = \pi(0,0,1) \cdot \nu \\ \Leftrightarrow \quad & \lambda + \gamma = \frac{\lambda + \gamma}{\nu} \cdot \nu \quad \Leftrightarrow \quad \lambda + \gamma = \lambda + \gamma. \end{aligned}$$

For $n > 0$ and $k = 0$ holds

$$\begin{aligned} &\pi(n,0,1) \cdot \nu = \pi(n+1,1,0) \cdot \mu \\ \Leftrightarrow \quad & \left(\frac{\lambda}{\mu}\right)^n \cdot \frac{\lambda}{\nu} \cdot \nu = \left(\frac{\lambda}{\mu}\right)^{n+1} \cdot \mu \quad \Leftrightarrow \quad \frac{\lambda^{n+1}}{\mu^n} = \frac{\lambda^{n+1}}{\mu^n}. \end{aligned}$$

For $n > 0$ and $k = 1$ holds

$$\begin{aligned} &\pi(n,1,0) \cdot \Big(\lambda + \mu\Big) = \pi(n-1,1,0) \cdot \lambda + \pi(n,0,1) \cdot \nu \\ \Leftrightarrow \quad & \left(\frac{\lambda}{\mu}\right)^n \cdot \Big(\lambda + \mu\Big) = \left(\frac{\lambda}{\mu}\right)^{n-1} \cdot \lambda + \left(\frac{\lambda}{\mu}\right)^n \cdot \frac{\lambda}{\nu} \cdot \nu \\ \Leftrightarrow \quad & \frac{\lambda^{n+1}}{\mu^n} + \frac{\lambda^n}{\mu^{n-1}} = \frac{\lambda^n}{\mu^{n-1}} + \frac{\lambda^{n+1}}{\mu^n}. \end{aligned}$$

C can be calculated by the normalizing condition. Hence,

$$\begin{aligned} C &= \frac{\lambda+\gamma}{\nu} + \frac{\lambda}{\nu} \cdot \left(\sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n - 1 \right) + \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n = \frac{\lambda+\gamma}{\nu} + \frac{\lambda}{\nu} \cdot \left(\frac{1}{1-\frac{\lambda}{\mu}} - 1 \right) + \frac{1}{1-\frac{\lambda}{\mu}} \\ &= \frac{\lambda+\gamma}{\nu} + \frac{\lambda}{\nu} \cdot \left(\frac{\mu}{\mu-\lambda} - 1 \right) + \frac{\mu}{\mu-\lambda} = \frac{\lambda}{\nu} + \frac{\gamma}{\nu} + \frac{\lambda}{\nu} \cdot \left(\frac{\lambda}{\mu-\lambda} \right) + \frac{\mu}{\mu-\lambda} \\ &= \frac{\gamma}{\nu} + \frac{\lambda}{\nu} \cdot \left(1 + \frac{\lambda}{\mu-\lambda} \right) + \frac{\mu}{\mu-\lambda} = \frac{\gamma}{\nu} + \frac{\lambda}{\nu} \cdot \left(\frac{\mu}{\mu-\lambda} \right) + \frac{\mu}{\mu-\lambda} \\ &= \frac{\gamma}{\nu} + \frac{\mu}{\mu-\lambda} \cdot \left(1 + \frac{\lambda}{\nu} \right). \end{aligned}$$

If we rewrite the limiting and stationary distribution as follows

$$\begin{aligned} \pi(0,0,1) &= C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^0 \cdot \frac{\lambda+\gamma}{\nu}, & \\ \pi(n,0,1) &= C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n \cdot \frac{\lambda}{\nu}, & n > 0, \\ \pi(0,1,0) &= C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^0 \cdot 1, & \\ \pi(n,1,0) &= C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n \cdot 1, & n > 0, \end{aligned}$$

we see that in fact the limiting and stationary distribution is not of product form with $\xi(n) = C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n$, $n \in \mathbb{N}_0$, since $\theta(k, b-k)$ cannot be defined independently of n as we argued in Proposition 4.2.35. $\square$

Proposition 4.2.37. *The throughput TH of the queueing-inventory system with base stock level $b = 1$ and queue-length-independent service rates μ and arrival rates λ is*

$$TH = \frac{\lambda\nu}{\nu + \lambda + \gamma - \frac{\lambda\gamma}{\mu}}.$$

Proof. The throughput can be calculated as follows

$$\begin{aligned}
TH &= \mu \cdot \sum_{n=1}^{\infty} \pi(n,1,0) \overset{(4.2.15)}{=} \mu \cdot \sum_{n=1}^{\infty} C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n \overset{(4.2.16)}{=} \mu \cdot \frac{\frac{\lambda}{\mu}}{C \cdot \left(1 - \frac{\lambda}{\mu}\right)} \\
&= \frac{\lambda}{\left(\frac{\mu}{\mu-\lambda} \cdot \left(1 + \frac{\lambda}{\nu}\right) + \frac{\gamma}{\nu}\right) \cdot \frac{\mu-\lambda}{\mu}} = \frac{\lambda}{\left(1 + \frac{\lambda}{\nu}\right) + \frac{\gamma}{\nu} \cdot \frac{\mu-\lambda}{\mu}} = \frac{\lambda}{\frac{\nu+\lambda}{\nu} + \frac{\gamma}{\nu} \cdot \frac{\mu-\lambda}{\mu}} \\
&= \frac{\lambda\nu}{\nu + \lambda + \gamma - \frac{\gamma\lambda}{\mu}}.
\end{aligned}$$

□

4.3. Separable systems: Multiple locations

As we have seen in Section 4.2.3 closed form expressions are not available for the queueing-inventory system with $b \geq 2$. The question is: "Can we modify the queueing-inventory system so that we get product form results?" This is even true for a supply chain with $J > 1$ locations as we see in this section.

4.3.1. Description of the general model

The supply chain of interest is depicted in Figure 4.3.1. We have a set of locations $\overline{J} := \{1, 2, \ldots, J\}$. Each of the locations consists of a production system with an attached inventory with perishable raw material. The inventories are replenished by a single central supplier, which is referred to as workstation $J+1$ and manufactures raw material for all locations. The items of raw material are indistinguishable (exchangeable).

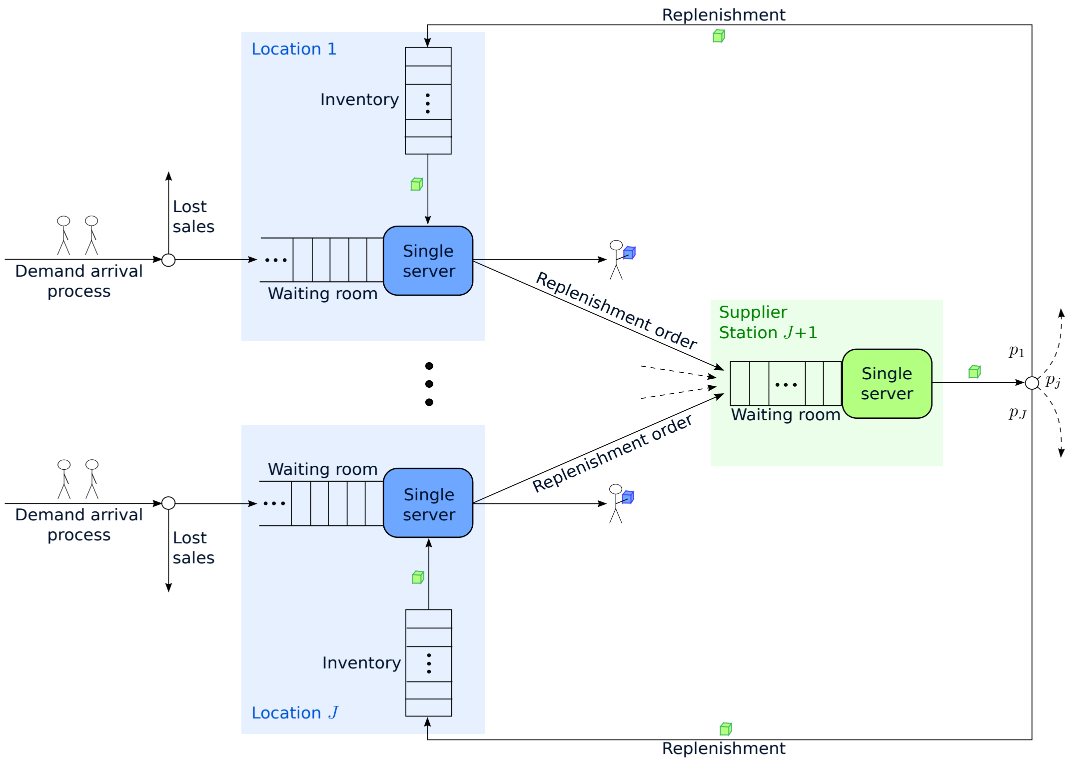


Figure 4.3.1.: Supply chain with base stock policy

Facilities in the supply chain. Each production system $j \in \overline{J}$ consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under a FCFS regime. Customers arrive one by one at production system j according to a Poisson process with rate $\lambda_j > 0$ and require service. To satisfy a customer's demand the production system needs exactly one item of raw material, which is taken from the associated local inventory. When a new customer arrives at a location while the previous customer's order is not finished, this customer will wait. If the inventory is depleted at location j, the customers who are already waiting in line will wait, but new arriving customers at this location will decide not to join the queue and are lost ("local lost sales").

The service requests at the locations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at location j is provided with local queue-length-dependent intensity. If there are $n_j > 0$ customers present at location j either waiting or in service (if any) and if the inventory is not depleted, the service intensity is $\mu_j(n_j) > 0$. If the server is ready to serve a customer who is at the head of the line, and the inventory is not depleted, the service immediately starts. Otherwise, the service starts at the instant of time when the next replenishment arrives at the local inventory.

For the control of the inventories we use base stock policies. Thus, each item taken from the inventory results in a direct order for one item of raw material. This means, if a served customer departs from the system or a raw material is perished, an order of the consumed resp. of the perished raw material is placed at the central supplier at this time instant. The local base stock level $b_j \geq 1$ is the maximal size of the inventory at location j, we denote $\mathbf{b} := \big(b_j : j \in \overline{J}\big)$. Note that there can be more than one outstanding order.

The items of raw material in the inventories are perishable. In the previous section, we take into consideration whether there are customers in the system. However, closed form expressions were not available. In this section, we look at a modification:
If there are $k_j > 0$ items of raw material in the inventory at location $j \in \overline{J}$, the loss rates of inventory due to perishing are $\gamma_j \cdot d_j\,(k_j)$. We call the functions $k_j \mapsto \gamma_j \cdot d_j\,(k_j)$ ageing regimes.

The central supplier consists of a single server (machine) and a waiting room under FCFS regime. At most $\sum_{j\in\overline{J}} b_j - 1$ replenishment orders are waiting at the central supplier. Service times at the central supplier are exponentially distributed with parameter $\nu > 0$.

Routing in the supply chain. A customer departs from the system immediately after the service and the associated consumed raw material is removed from the inventory at this time instant.

A finished item of raw material departs immediately from the central supplier and is sent to location $j \in \overline{J}$ with probability $p_j > 0$, independent of the network's history. A predetermined delivering schedule is represented by $(p_j : j \in \overline{J})$ with $\sum_{j\in\overline{J}} p_j = 1$. If the inventory is not full at location j (this means that the on-hand inventory level at location j is lower than the base stock level b_j), the item is added to the inventory at this location. Otherwise the central supplier will spend extra time on the already finished item of raw material and resend it to a new location $i \in \overline{J}$ according to the predetermined probabilities p_i, independent of the network's history. We remark that this is equivalent to the assumption that the item of raw material is discarded and a new replenishment order is added at the central supplier.

We assume that the replenished raw materials are "fresh". Much of the literature on perishable items assumes this to avoid to complicate the model (cf. [Bar10, p. 2]).

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the central supplier and the local inventories are negligible.

All inter-arrival times, service times and life times of items constitute an independent family of random variables.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the size of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_{J+1}(t)$ we denote the number of replenishment orders at the central supplier at time $t \geq 0$ either waiting or in service (queue length).

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), \ldots, X_J(t), Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) \,:\, t \geq 0) \,.$$

Then, due to the usual independence assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \left\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J,\, \mathbf{k} \in K\right\}$$

with

$$K := \{(k_1, \ldots, k_J, k_{J+1}) | 0 \leq k_j \leq b_j,\ j = 1, \ldots, J,\ k_{J+1} = \sum_{j=1}^{J} (b_j - k_j)\} \subset \mathbb{N}_0^{J+1}.$$

4.3.2. Limiting and stationary distribution

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z; \tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n}, \mathbf{k}) \in E$:

$$\begin{aligned}
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n} + \mathbf{e}_i, \mathbf{k})\right) &= \lambda_i \cdot 1_{\{k_i > 0\}}, && i \in \overline{J}, \\
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n}, \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1})\right) &= \gamma_i \cdot d_i\left(k_i\right) \cdot 1_{\{k_i > 0\}}, && i \in \overline{J}, \\
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n} - \mathbf{e}_i, \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1})\right) &= \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}}, && i \in \overline{J}, \\
q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n}, \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1})\right) &= \nu\, p_i \cdot 1_{\{k_i < b_i\}}, && i \in \overline{J}.
\end{aligned}$$

Furthermore, $q(z; \tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z; z\right) = -\sum_{\substack{\tilde{z} \in E, \\ z \neq \tilde{z}}} q\left(z; \tilde{z}\right) \qquad \forall z \in E.$$

Note that $k_{J+1} > 0$ holds if $k_i < b_i$ for some $i \in \overline{J}$.

Proposition 4.3.1. *The strictly positive measure* $\mathbf{x} := (x\left(\mathbf{n}, \mathbf{k}\right) : (\mathbf{n}, \mathbf{k}) \in E)$ *with*

$$x\left(\mathbf{n}, \mathbf{k}\right) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}\left(\mathbf{k}\right), \tag{4.3.1}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{4.3.2}$$

$$\widetilde{\theta}(\mathbf{k}) = \widetilde{\theta}(k_1, \dots, k_J, k_{J+1}) = \prod_{j \in \overline{J}} \prod_{\ell=1}^{k_j} \frac{\nu \, p_j}{\lambda_j + \gamma_j \cdot d_j(\ell)}, \quad \mathbf{k} \in K, \tag{4.3.3}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z*.*

Proof. Therefore, the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are for $(\mathbf{n}, \mathbf{k}) \in E$ given by

$$\begin{aligned}
& x(\mathbf{n}, \mathbf{k}) \\
& \cdot \Big(\sum_{i \in \overline{J}} (\lambda_i + \gamma_i \cdot d_i(k_i)) \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \, p_i \cdot 1_{\{k_i < b_i\}} \Big) \\
= & \sum_{i \in \overline{J}} x(\mathbf{n} - \mathbf{e}_i, \mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} \\
& + \sum_{i \in \overline{J}} x(\mathbf{n} + \mathbf{e}_i, \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \mu_i(n_i + 1) \cdot 1_{\{k_i < b_i\}} \\
& + \sum_{i \in \overline{J}} x(\mathbf{n}, \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \gamma_i \cdot d_i(k_i + 1) \cdot 1_{\{k_i < b_i\}} \\
& + \sum_{i \in \overline{J}} x(\mathbf{n}, \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \, p_i \cdot 1_{\{k_i > 0\}}.
\end{aligned}$$

It has to be shown that the stationary measure (4.3.1) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

Substitution of (4.3.1) and (4.3.2) into the global balance equations directly leads to

$$\begin{aligned}
& \Big(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \Big) \cdot \widetilde{\theta}(\mathbf{k}) \\
& \cdot \Big(\sum_{i \in \overline{J}} (\lambda_i + \gamma_i \cdot d_i(k_i)) \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \, p_i \cdot 1_{\{k_i < b_i\}} \Big) \\
= & \sum_{i \in \overline{J}} \Big(\prod_{j \in \overline{J} \setminus \{i\}} \widetilde{\xi}_j(n_j) \Big) \cdot \widetilde{\xi}_i(n_i - 1) \cdot \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} \\
& + \sum_{i \in \overline{J}} \Big(\prod_{j \in \overline{J} \setminus \{i\}} \widetilde{\xi}_j(n_j) \Big) \cdot \widetilde{\xi}_i(n_i + 1) \cdot \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \mu_i(n_i + 1) \cdot 1_{\{k_i < b_i\}} \\
& + \sum_{i \in \overline{J}} \Big(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \Big) \cdot \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \gamma_i \cdot d_i(k_i + 1) \cdot 1_{\{k_i < b_i\}} \\
& + \sum_{i \in \overline{J}} \Big(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \Big) \cdot \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \, p_i \cdot 1_{\{k_i > 0\}}.
\end{aligned}$$

By substitution of (4.3.2) we obtain

$$
\begin{aligned}
&\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k})\\
&\cdot\Big(\sum_{i\in\overline{J}}(\lambda_i+\gamma_i\cdot d_i\,(k_i))\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\nu\,p_i\cdot 1_{\{k_i<b_i\}}\Big)\\
=&\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k})\cdot\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{k_i>0\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot\lambda_i\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot\gamma_i\cdot d_i\,(k_i+1)\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\,(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\cdot\nu\,p_i\cdot 1_{\{k_i>0\}}.
\end{aligned}
$$

Cancelling $\left(\prod_{j\in\overline{J}}\xi_j(n_j)\right)$ and the sum with the terms $\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{k_i>0\}}$ on both sides of the equation leads to

$$
\begin{aligned}
&\widetilde{\theta}\,(\mathbf{k})\Big(\sum_{i\in\overline{J}}(\lambda_i+\gamma_i\cdot d_i\,(k_i))\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\nu\,p_i\cdot 1_{\{k_i<b_i\}}\Big)\\
=&\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot\lambda_i\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot\gamma_i\cdot d_i\,(k_i+1)\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\cdot\nu\,p_i\cdot 1_{\{k_i>0\}}.
\end{aligned}
$$

This is equivalent to

$$
\begin{aligned}
&\widetilde{\theta}\,(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}(\lambda_i+\gamma_i\cdot d_i\,(k_i))\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\nu\,p_i\cdot 1_{\{k_i<b_i\}}\Big)\\
=&\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}+\mathbf{e}_i-\mathbf{e}_{J+1})\cdot(\lambda_i+\gamma_i\cdot d_i\,(k_i+1))\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\widetilde{\theta}\,(\mathbf{k}-\mathbf{e}_i+\mathbf{e}_{J+1})\cdot\nu\,p_i\cdot 1_{\{k_i>0\}}.
\end{aligned}
\tag{4.3.4}
$$

Substitution of (4.3.3) leads to

$$\begin{aligned}
&\widetilde{\theta}(\mathbf{k})\Big(\sum_{i\in\overline{J}}(\lambda_i+\gamma_i\cdot d_i(k_i))\cdot 1_{\{k_i>0\}}+\sum_{i\in\overline{J}}\nu\, p_i\cdot 1_{\{k_i<b_i\}}\Big)\\
=&\sum_{i\in\overline{J}}\widetilde{\theta}(\mathbf{k})\cdot\frac{\nu\, p_i}{\lambda_i+\gamma_i\cdot d_i(k_i+1)}\cdot(\lambda_i+\gamma_i\cdot d_i(k_i+1))\cdot 1_{\{k_i<b_i\}}\\
&+\sum_{i\in\overline{J}}\widetilde{\theta}(\mathbf{k})\cdot\frac{\lambda_i+\gamma_i\cdot d_i(k_i)}{\nu\, p_i}\cdot\nu\, p_i\cdot 1_{\{k_i>0\}}.
\end{aligned}$$

The right-hand side of the last equation is

$$\sum_{i\in\overline{J}}\nu\, p_i\cdot 1_{\{k_i<b_i\}}+\sum_{i\in\overline{J}}(\lambda_i+\gamma_i\cdot d_i(k_i))\cdot 1_{\{k_i>0\}},$$

which is obviously the left-hand side.

Inspection of the system (4.3.4) reveals that it is a “generator equation”, i.e. the global balance equation $\widetilde{\theta}\cdot\mathbf{Q}_{red}=\mathbf{0}$ for a suitably defined ergodic Markov process on state space K with “reduced generator” $\mathbf{Q}_{red}$. Because the Markov process generated by $\mathbf{Q}_{red}$ is irreducible the solution of (4.3.4) is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Remark 4.3.2. $\widetilde{\theta}(\mathbf{k})=\widetilde{\theta}(k_1,\ldots,k_J,k_{J+1})$ is obtained as a strictly positive solution of (4.3.4) which resembles the global balance equations of an artificial non-standard Gordon-Newell network of queues with $J+1$ nodes and $\sum_{j\in\overline{J}}b_j$ customers, exponentially distributed service times with state-dependent rates $(\lambda_j+\gamma_j\cdot d_j(k_j))$ for $k_j\le b_j$ and “∞” otherwise at node $j\in\{1,\ldots,J\}$ and with rate ν at node $J+1$ and state-dependent routing probabilities. More precisely, it is a starlike system with $r(j,J+1)=1$, $j\in\overline{J}$, and branching probabilities $r(J+1,j)=p_j$, $j\in\overline{J}$.

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 4.3.3. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi:=(\pi(\mathbf{n},\mathbf{k}):(\mathbf{n},\mathbf{k})\in E),\quad \pi(\mathbf{n},\mathbf{k}):=\lim_{t\to\infty}P(Z(t)=(\mathbf{n},\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi:=\left(\xi(\mathbf{n}):\mathbf{n}\in\mathbb{N}_0^J\right),\quad \xi(\mathbf{n}):=\lim_{t\to\infty}P((X_1(t),\ldots,X_J(t))=\mathbf{n}),$$

$$\theta:=(\theta(\mathbf{k}):\mathbf{k}\in K),\quad \theta(\mathbf{k}):=\lim_{t\to\infty}P((Y_1(t),\ldots,Y_J(t),W_{J+1}(t))=\mathbf{k}).$$

Theorem 4.3.4. *The queueing-inventory process Z is ergodic if and only if for $j\in\overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}<\infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n},\mathbf{k}) = \xi(\mathbf{n})\cdot\theta(\mathbf{k}), \tag{4.3.5}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1}\cdot\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}, \quad n_j\in\mathbb{N}_0,\ j\in\overline{J}, \tag{4.3.6}$$

$$\theta(\mathbf{k}) = \theta(k_1,\dots,k_J,k_{J+1}) = C_\theta^{-1}\cdot\prod_{j\in\overline{J}}\prod_{\ell=1}^{k_j}\frac{\nu\,p_j}{\lambda_j+\gamma_j\cdot d_j(\ell)}, \quad \mathbf{k}\in K, \tag{4.3.7}$$

and normalisation constants

$$C_j = \sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)} \qquad \text{and} \qquad C_\theta = \sum_{\mathbf{k}\in K}\prod_{j\in\overline{J}}\prod_{\ell=1}^{k_j}\frac{\nu\,p_j}{\lambda_j+\gamma_j\cdot d_j(\ell)}.$$

Proof. Z is ergodic, if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x}\cdot\mathbf{Q}=\mathbf{0}$ from Proposition 4.3.1 can be normalised (i.e. $\sum_{\mathbf{n}\in\mathbb{N}_0}\sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k}) < \infty$). Because of Proposition 4.3.1 it holds

$$\begin{aligned}\sum_{\mathbf{n}\in\mathbb{N}_0}\sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k}) &= \sum_{\mathbf{n}\in\mathbb{N}_0}\widetilde{\xi}(\mathbf{n})\cdot\sum_{\mathbf{k}\in K}\widetilde{\theta}(\mathbf{k})\\ &= \left(\prod_{j\in\overline{J}}\sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}\right)\cdot\left(\sum_{\mathbf{k}\in K}\prod_{j\in\overline{J}}\prod_{\ell=1}^{k_j}\frac{\nu\,p_j}{\lambda_j+\gamma_j\cdot d_j(\ell)}\right).\end{aligned}$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 4.3.1 can be normalised if and only if $\sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)} < \infty$ for all $j\in\overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n},\mathbf{k}) = \frac{x(\mathbf{n},\mathbf{k})}{\sum_{\mathbf{n}\in\mathbb{N}_0}\sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k})},$$

where $x(\mathbf{n},\mathbf{k})$ is given in Proposition 4.3.1. □

Remark 4.3.5. It has to be noted that k_{J+1} occurs only implicitly on the right side of (4.3.7). This hides a strong negative correlation of the coordinate processes $W_{J+1}(t)$ and $(Y_1(t),\dots,Y_J(t))$ which is due to the state space restrictions.

The expression (4.3.5) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

Representation (4.3.6) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}}\xi_j(n_j) = \prod_{j\in\overline{J}} C_j^{-1}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. However, in our production network servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (4.3.5) has been unexpected to us.

Our production-inventory system can be considered as a "Jackson network in a random environment" in [KDO16, Section 4]. We can interpret the inventory-replenishment subsystem, which contributes via θ to (4.3.5), as a "random environment" for the production network of nodes $\overline{J}$, which is a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1] we conclude from the hindsight that decoupling of the queueing process $(X_1, \ldots, X_J)$ and the process $(Y_1, \ldots, Y_J, W_{J+1})$, i.e. the formula (4.3.5), is a consequence of that Theorem 4.1.

Our direct proof of Theorem 4.3.4 is much shorter than embedding the present model into the general framework of [KDO16].

The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

4.3.3. Separable approximation of non-separable systems

As we have seen in Section 4.2.3 closed form expressions are not available for the single location production-inventory system with perishable items where ageing is dependent on whether an item from inventory is already in usage by production for base stock levels $b \geq 2$. The question is: "Can we use the product form results for a single location from Section 4.3.2 to obtain simple product form bounds for the system with unknown non-product form stationary distribution in Section 4.2.3?"

Motivated by the works of van Dijk and this coauthors (e.g. [Dij11b, Section 1,7, pp. 62f.], [Dij98, pp. 311ff.], [DK92], [DW89]) we are interested in lower and upper product form bounds for the throughput of the production-inventory system with unknown non-product form stationary distribution. Van Dijk shows in [Dij11b, Section 1,7, pp. 62f.] for the simple but unsolvable ($\equiv$ unknown stationary distribution) tandem queue with both finite first and finite second station, that the product form (modification) turns out to be quite fruitful to provide a simple (lower and upper) product form bound for the throughput.

It can intuitively be expected that bounds for the non-product production-inventory system from Section 4.2.3 can be built by single location production-inventory systems[4] from Section 4.3.2 ($\lambda_1 := \lambda$, $\mu_1 := \mu$, $p_1 := 1$ and $b_1 := b$) with perishable items where ageing is independent of whether an item from inventory is already in usage by production. The intuitive explanation of this is as follows: If we have either more inventory and at least the same number of customers in system, or more customers in system and at least the same stock size of the inventory, the system should be able to produce more output.

A lower bound is built by the production-inventory system with perishable items where all items in the inventory are subject to ageing — even the one already reserved for production. The ageing regime in state $(m, k) \in E$ is $k \mapsto \gamma_1 \cdot d_1(k) := \gamma \cdot k$. This production-inventory system is called "$-$"-system. We denote henceforth the state process of this system by Z^-, the stationary distribution of this system under ergodicity by $\pi^- = (\pi^-(m, k) : (m, k) \in E)$ and the throughput for the production-inventory system with this ageing regime by

$$TH^- = \sum_{(m,k)\in E} \pi^-(m, k) \cdot \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}}. \tag{4.3.8}$$

An upper bound is built by the production-inventory system with perishable items where one item in the inventory (if there is any) is not subject to ageing — even if the server is idling and no item is reserved for production. This results in an ageing regime in state $(m, k) \in E$ of $k \mapsto \gamma_1 \cdot d_1(k) := \gamma \cdot (k - 1)_+$. This production-inventory system is called "+"-system. We denote henceforth the state process of this system by Z^+, the stationary distribution of this system under ergodicity by $\pi^+ = (\pi^+(m, k) : (m, k) \in E)$ and the throughput for the production-inventory system with this ageing regime by

$$TH^+ = \sum_{(m,k)\in E} \pi^+(m, k) \cdot \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}}. \tag{4.3.9}$$

[4] We will henceforth use an abbreviated notation because the base stock level is a fixed parameter: $(n, k) := (n, k, b - k)$.

The production-inventory system with unknown non-product form stationary distribution from Section 4.2.3 with perishable items where ageing is dependent on whether an item from inventory is already in usage by production (i.e. the server is busy, or if the server is idling and the inventory is not empty) henceforth will be called "o"-system. The ageing regime in state $(m,k) \in E$ is $k \mapsto (\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{m>0\}}$. We denote the state process of this system by Z^o henceforth. In this case, the throughput for the ergodic production-inventory system with perishable items is with stationary distribution $\pi^o = (\pi^o(m,k) : (m,k) \in E)$

$$TH^o = \sum_{(m,k)\in E} \pi^o(m,k) \cdot \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}}. \tag{4.3.10}$$

The ageing regimes are ordered in the following way. For $(m,k) \in E$ holds

$$\gamma \cdot (k-1)_+ \leq (\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{m>0\}} \leq \gamma \cdot k.$$

This leads to the following conjecture.

Conjecture 4.3.6 (Monotonicity of throughputs). *Consider three (exponential) ergodic production-inventory systems with the same arrival rate λ, service rate μ, replenishment rate ν, individual ageing rate γ for items in the inventory which are subject to ageing.*

Then the following monotonicity property for the throughputs holds

$$TH^- \leq TH^o \leq TH^+.$$

This conjecture is supported by the following results in Section 4.3.3.1 for systems with base stock level $b = 1$ and in Section 4.3.3.2 for special systems with base stock level $b \geq 2$.

4.3.3.1. Production-inventory system with base stock level $b = 1$

Firstly, we will explicitly compute the throughputs of the systems since we have closed form expressions for the stationary distributions of the "o"-system as well as of the "$-$"-system and "+"-system for the case of base stock level equal to one.

The throughput TH of the "o"-system with base stock level $b = 1$ was calculated in Proposition 4.2.37 and is

$$TH^o = \frac{\lambda \cdot \nu}{\lambda + \nu + \gamma - \frac{\lambda \cdot \gamma}{\mu}}.$$

Proposition 4.3.7. *The throughput of the "$-$"-system with base stock level $b = 1$ is*

$$TH^- = \frac{\lambda \cdot \nu}{\lambda + \gamma + \nu}.$$

Proof. From Theorem 4.3.4 we know that the limiting and stationary distribution of the "$-$"-system is given by

$$\pi^-(n,k) = \xi^-(n) \cdot \theta^-(k), \tag{4.3.11}$$

with

$$\xi^-(n) = \left(C^-\right)^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n, \quad n \in \mathbb{N}_0, \tag{4.3.12}$$

$$\theta^{-}(k) = \left(C_{\theta}^{-}\right)^{-1} \cdot \prod_{\ell=1}^{k} \left(\frac{\nu}{\lambda + \gamma \cdot \ell}\right), \quad k \in \{0,1\}, \tag{4.3.13}$$

and normalisation constants

$$\left(C^{-}\right)^{-1} = \left(1 - \frac{\lambda}{\mu}\right),$$
$$\left(C_{\theta}^{-}\right)^{-1} = \left(1 + \frac{\nu}{\lambda + \gamma}\right)^{-1} = \left(\frac{\lambda + \gamma + \nu}{\lambda + \gamma}\right)^{-1} = \left(\frac{\lambda + \gamma}{\lambda + \gamma + \nu}\right).$$

The throughput TH^{-} of the "−"-system can be calculated as follows

$$\begin{aligned} TH^{-} &= \mu \cdot \sum_{n=1}^{\infty} \pi^{-}(n,1) = \mu \cdot \sum_{n=1}^{\infty} \xi^{-}(n) \cdot \theta^{-}(1) \\ &= \mu \cdot \sum_{n=1}^{\infty} \left(1 - \frac{\lambda}{\mu}\right) \cdot \left(\frac{\lambda}{\mu}\right)^{n} \cdot \left(\frac{\lambda + \gamma}{\lambda + \gamma + \nu}\right) \cdot \left(\frac{\nu}{\lambda + \gamma}\right) = \frac{\lambda \cdot \nu}{\lambda + \gamma + \nu}. \end{aligned}$$

□

Proposition 4.3.8. *The throughput for the "+"-system with base stock level $b = 1$ is*

$$TH^{+} = \frac{\lambda \cdot \nu}{\lambda + \nu}.$$

Proof. From Theorem 4.3.4 we know that the limiting and stationary system of the "+"-system is given by

$$\pi^{+}(n,k) = \xi^{+}(n) \cdot \theta^{+}(k), \tag{4.3.14}$$

with

$$\xi^{+}(n) = \left(C^{+}\right)^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^{n}, \quad n \in \mathbb{N}_0, \tag{4.3.15}$$

$$\theta^{+}(k) = \left(C_{\theta}^{+}\right)^{-1} \prod_{\ell=1}^{k} \left(\frac{\nu}{\lambda + \gamma \cdot (\ell - 1)}\right), \quad k \in \{0,1\}, \tag{4.3.16}$$

and normalisation constants

$$\left(C^{+}\right)^{-1} = \left(1 - \frac{\lambda}{\mu}\right) \quad \text{and} \quad \left(C_{\theta}^{+}\right)^{-1} = \left(1 + \frac{\nu}{\lambda}\right)^{-1} = \left(\frac{\lambda + \nu}{\lambda}\right)^{-1} = \left(\frac{\lambda}{\lambda + \nu}\right).$$

The throughput TH^{+} of the "+"-system can be calculated as follows

$$\begin{aligned} TH^{+} &= \mu \cdot \sum_{n=1}^{\infty} \pi^{+}(n,1) = \mu \cdot \sum_{n=1}^{\infty} \xi^{+}(n) \cdot \theta^{+}(1) \\ &= \mu \cdot \sum_{n=1}^{\infty} \left(1 - \frac{\lambda}{\mu}\right) \cdot \left(\frac{\lambda}{\mu}\right)^{n} \cdot \left(\frac{\nu}{\lambda + \gamma}\right) \cdot \left(\frac{\lambda}{\lambda + \nu}\right) = \frac{\lambda \cdot \nu}{\lambda + \nu}. \end{aligned}$$

□

Corollary 4.3.9 (Monotonicity of throughputs). *Consider three (exponential) ergodic production-inventory systems with the same arrival rate λ, service rate μ, replenishment rate ν, individual ageing rate γ for items in the inventory which are subject to ageing and base stock level $b = 1$.*

Then the following monotonicity property for the throughputs holds

$$TH^- < TH < TH^+.$$

4.3.3.2. Production-inventory system with base stock level $b \geq 2$

We assume henceforth in this section $b \geq 2$.

Preliminaries

Queue-length-dependent ageing

We consider perishable items where ageing is dependent on whether an item from inventory is already in usage by production, i.e. the server is busy, or if the server is idling and the inventory is not empty. The loss rate of inventory due to perishing in state $(m, k) \in E$ is

$$(\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{m>0\}}.$$

Z^o is uniformizable with constant $\alpha := \lambda + \mu + \nu + \gamma \cdot b$. So, whenever the underlying Poisson-α process indicates a jump, the state of the associated uniformization chain Z_u^o jumps with probability 1. Z_u^o is ergodic with stationary distribution π^o as well. We denote the one-step transition probability of Z_u^o by $R^o = (R^o(z; \widetilde{z}) : z, \widetilde{z} \in E)$, and have for $k = 0$ and $n \geq 0$

$$R^o((n,0);(n,1)) = \frac{\nu}{\alpha}, \qquad R^o((n,0);(n,0)) = \frac{\lambda + \mu + \gamma \cdot b}{\alpha},$$

for $k \in \{1, \dots, b-1\}$ and $n = 0$

$$R^o((0,k);(1,k)) = \frac{\lambda}{\alpha}, \qquad R^o((0,k);(0,k+1)) = \frac{\nu}{\alpha},$$
$$R^o((0,k);(0,k-1)) = \frac{\gamma \cdot k}{\alpha}, \qquad R^o((0,k);(0,k)) = \frac{\mu + \gamma \cdot (b-k)}{\alpha},$$

for $k \in \{1, \dots, b-1\}$ and $n \geq 1$

$$R^o((n,k);(n+1,k)) = \frac{\lambda}{\alpha}, \qquad R^o((n,k);(n-1,k-1)) = \frac{\mu}{\alpha},$$
$$R^o((n,k);(n,k+1)) = \frac{\nu}{\alpha}, \qquad R^o((n,k);(n,k-1)) = \frac{\gamma \cdot (k-1)}{\alpha},$$
$$R^o((n,k);(n,k)) = \frac{\gamma \cdot (b-k+1)}{\alpha},$$

for $k = b$ and $n = 0$

$$R^o((0,b);(1,b)) = \frac{\lambda}{\alpha}, \qquad R^o((0,b);(0,b-1)) = \frac{\gamma \cdot b}{\alpha},$$
$$R^o((0,b);(0,b)) = \frac{\nu + \mu}{\alpha},$$

for $k = b$ and $n \geq 1$

$$R^o((n,b)\,;(n+1,b)) = \frac{\lambda}{\alpha}, \qquad R^o((n,b)\,;(n-1,b-1)) = \frac{\mu}{\alpha},$$
$$R^o((n,b)\,;(n,b-1)) = \frac{\gamma \cdot (b-1)}{\alpha}, \qquad R^o((n,b)\,;(n,b)) = \frac{\nu+\gamma}{\alpha}.$$

We associate with the uniformization chain Z_u^o a reward chain with one-step immediate reward vector

$$r = (r(m,k) : (m,k) \in E), \quad \text{with} \quad r(m,k) = \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}}, \tag{4.3.17}$$

so

$$TH^o = \sum_{(m,k)\in E} \pi^o(m,k) \cdot r(m,k). \tag{4.3.18}$$

We define the n-period reward vector of the reward chain as

$$v_n^o = (v_n^o(m,k) : (m,k) \in E), \quad n \geq 1, \tag{4.3.19}$$

which is according to van der Wal [vdW89, Lemma 2]

$$v_n^o = \sum_{h=0}^{n-1} (R^o)^h \cdot r, \quad n \geq 1, \tag{4.3.20}$$

and

$$v_{n+1}^o = r + R^o \cdot v_n^o, \quad n \geq 1. \tag{4.3.21}$$

From ergodicity of Z_u^o and the equality of the stationary distributions of Z^o and Z_u^o it follows that for any initial state $(m,k) \in E$ holds (cf. [vdW89, Lemma 2])

$$TH^o = \lim_{n\to\infty} \frac{1}{n} \cdot v_n^o(m,k). \tag{4.3.22}$$

Queue-length-independent ageing

We next consider perishable items where ageing is independent of whether an item from inventory is already in usage by production. We distinguish two modes of ageing.

Under the ageing regime for the lower bound, all items in the inventory are subject to ageing — even the one already reserved for production. The loss rate of inventory due to perishing in state $(m,k) \in E$ is $\gamma \cdot k$.

The state process Z^- of this system is uniformizable with constant $\alpha := \lambda+\mu+\nu+\gamma\cdot b$. The uniformization chain Z_u^- is ergodic with stationary distribution π^- as well. We denote the one-step transition probability of Z_u^- by $R^- = (R^-(z;\widetilde{z}) :\ z,\widetilde{z} \in E)$, and have for $k = 0$ and $n \geq 0$

$$R^-((n,0)\,;(n,1)) = \frac{\nu}{\alpha}, \qquad R^-((n,0)\,;(n,0)) = \frac{\lambda+\mu+\gamma\cdot b}{\alpha},$$

for $k \in \{1,\ldots,b-1\}$ and $n = 0$

$$R^-((0,k)\,;(1,k)) = \frac{\lambda}{\alpha}, \qquad R^-((0,k)\,;(0,k+1)) = \frac{\nu}{\alpha},$$
$$R^-((0,k)\,;(0,k-1)) = \frac{\gamma\cdot k}{\alpha}, \qquad R^-((0,k)\,;(0,k)) = \frac{\mu+\gamma\cdot(b-k)}{\alpha},$$

for $k \in \{1, \dots, b-1\}$ and $n \geq 1$

$$R^-((n,k)\,;(n+1,k)) = \frac{\lambda}{\alpha}, \qquad R^-((n,k)\,;(n-1,k-1)) = \frac{\mu}{\alpha},$$

$$R^-((n,k)\,;(n,k+1)) = \frac{\nu}{\alpha}, \qquad R^-((n,k)\,;(n,k-1)) = \frac{\gamma \cdot k}{\alpha},$$

$$R^-((n,k)\,;(n,k)) = \frac{\gamma \cdot (b-k)}{\alpha},$$

for $k = b$ and $n = 0$

$$R^-((0,b)\,;(1,b)) = \frac{\lambda}{\alpha}, \qquad R^-((0,b)\,;(0,b-1)) = \frac{\gamma \cdot b}{\alpha},$$

$$R^-((0,b)\,;(0,b)) = \frac{\nu + \mu}{\alpha},$$

for $k = b$ and $n \geq 1$

$$R^-((n,b)\,;(n+1,b)) = \frac{\lambda}{\alpha}, \qquad R^-((n,b)\,;(n-1,b-1)) = \frac{\mu}{\alpha},$$

$$R^-((n,b)\,;(n,b-1)) = \frac{\gamma \cdot b}{\alpha}, \qquad R^-((n,b)\,;(n,b)) = \frac{\nu}{\alpha}.$$

Under the ageing regime for the upper bound, one item in the inventory (if there is any) is not subject to ageing — even if the server is idling and no item is reserved for production. This results in a loss rate of inventory due to perishing in state $(m,k) \in E$ of $\gamma \cdot (k-1)_+$.

The state process Z^+ of this system is uniformizable with constant $\alpha := \lambda+\mu+\nu+\gamma \cdot b$. The uniformization chain Z_u^+ is ergodic with stationary distribution π^+ as well. We denote the one-step transition probability of Z_u^+ by $R^+ = (R^+(z;\widetilde{z}) : \, z, \widetilde{z} \in E)$, and have for $k = 0$ and $n \geq 0$

$$R^+((n,0)\,;(n,1)) = \frac{\nu}{\alpha}, \qquad R^+((n,0)\,;(n,0)) = \frac{\lambda + \mu + \gamma \cdot b}{\alpha},$$

for $k \in \{1, \dots, b-1\}$ and $n = 0$

$$R^+((0,k)\,;(1,k)) = \frac{\lambda}{\alpha}, \qquad R^+((0,k)\,;(0,k+1)) = \frac{\nu}{\alpha},$$

$$R^+((0,k)\,;(0,k-1)) = \frac{\gamma \cdot (k-1)}{\alpha}, \qquad R^+((0,k)\,;(0,k)) = \frac{\mu + \gamma \cdot (b-k+1)}{\alpha},$$

for $k \in \{1, \dots, b-1\}$ and $n \geq 1$

$$R^+((n,k)\,;(n+1,k)) = \frac{\lambda}{\alpha}, \qquad R^+((n,k)\,;(n-1,k-1)) = \frac{\mu}{\alpha}$$

$$R^+((n,k)\,;(n,k+1)) = \frac{\nu}{\alpha}, \qquad R^+((n,k)\,;(n,k-1)) = \frac{\gamma \cdot (k-1)}{\alpha},$$

$$R^+((n,k)\,;(n,k)) = \frac{\gamma \cdot (b-k+1)}{\alpha},$$

for $k = b$ and $n = 0$

$$R^+((0,b)\,;(1,b)) = \frac{\lambda}{\alpha}, \qquad R^+((0,b)\,;(0,b-1)) = \frac{\gamma \cdot (b-1)}{\alpha},$$

$$R^+((0,b)\,;(0,b)) = \frac{\nu + \mu + \gamma}{\alpha},$$

for $k = b$ and $n \geq 1$

$$R^{+}((n,b)\,;(n+1,b)) = \frac{\lambda}{\alpha}, \qquad R^{+}((n,b)\,;(n,b)) = \frac{\nu+\gamma}{\alpha},$$
$$R^{+}((n,b)\,;(n-1,b-1)) = \frac{\mu}{\alpha}, \qquad R^{+}((n,b)\,;(n,b-1)) = \frac{\gamma\cdot(b-1)}{\alpha}.$$

We associate with both uniformization chains Z_u^- and Z_u^+ under different ageing regimes a reward chain with one-step immediate reward vector (the same for all three production-inventory systems)

$$r = (r(m,k) : (m,k) \in E) \quad \text{with} \quad r(m,k) = \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}}, \tag{4.3.23}$$

so

$$TH^- = \sum_{(m,k)\in E} \pi^-(m,k)\cdot r(m,k) \quad \text{and} \quad TH^+ = \sum_{(m,k)\in E} \pi^+(m,k)\cdot r(m,k). \tag{4.3.24}$$

We define the n-period reward vectors of these reward chains as

$$v_n^- = (v_n^-(m,k) : (m,k)\in E) \quad \text{resp.} \quad v_n^+ = (v_n^+(m,k) : (m,k)\in E), \quad n \geq 1, \tag{4.3.25}$$

which implies [vdW89, Lemma 2] similarly as above

$$v_n^- = \sum_{h=0}^{n-1}(R^-)^h \cdot r \quad \text{resp.} \quad v_n^+ = \sum_{h=0}^{n-1}(R^+)^h\cdot r, \quad n\geq 1, \tag{4.3.26}$$

and

$$v_{n+1}^- = r + R^-\cdot v_n^- \quad \text{resp.} \quad v_{n+1}^+ = r + R^+\cdot v_n^+, \quad n \geq 1. \tag{4.3.27}$$

From ergodicity of Z_u^- and the equality of the stationary distributions of Z^- and Z_u^- it follows that for any initial state $(m,k)\in E$ holds

$$TH^- = \lim_{n\to\infty}\frac{1}{n}\cdot v_n^-(m,k). \tag{4.3.28}$$

Similarly, from ergodicity of Z_u^+ and the equality of the stationary distributions of Z^+ and Z_u^+ it follows that for any initial state $(m,k)\in E$ holds

$$TH^+ = \lim_{n\to\infty}\frac{1}{n}\cdot v_n^+(m,k). \tag{4.3.29}$$

Recall, that TH^- and TH^+ can be explicitly computed (cf. Theorem 4.3.4).

Our comparison relies on the expressions (4.3.22), (4.3.28) and (4.3.29) for the throughputs and on additional conditions which will guarantee that for all $n\in\mathbb{N}$ holds

$$v_n^-(m,k) \leq v_n^o(m,k) \leq v_n^+(m,k) \qquad \text{for all initial states} \quad (m,k)\in E, \tag{4.3.30}$$

for the finite time cumulative (expected) rewards. It turns out that essential properties to prove (4.3.30) are internal monotonicities of the cumulative rewards. To shorten the presentation we use the symbol "$*$" to refer to all the symbols "$o,-,+$" in the respective expressions.

Definition 4.3.10. We say that v_n^* is isotone (with respect to the natural order on $\mathbb{N}_0 \times \{0, 1, \dots, b\}$) if

$$\forall (m,k),(m',k') \in E: \ \big[(m \leq m' \wedge k \leq k')\big] \text{ implies } \big[v_n^*(m,k) \leq v_n^*(m',k')\big]. \quad (4.3.31)$$

Proposition 4.3.11. *Consider three (exponential) ergodic queueing inventory systems with the same arrival rate λ, service rate μ, replenishment rate ν, individual ageing rate γ for items in the inventory which are subject to ageing.*
The ageing regimes of the systems are different, which results in different Markovian state processes which we denote by Z^o under ageing regime $k \mapsto (\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{m>0\}}$, Z^- under ageing regime $k \mapsto \gamma \cdot k$, and Z^+ under ageing regime $k \mapsto \gamma \cdot (k-1)_+$. Then the following holds.

(a) *If v_n^- is isotone, then for all $(m,k) \in E$ holds*

$$v_n^-(m,k) \leq v_n^o(m,k) \quad \forall n \in \mathbb{N},$$

and consequently $TH^- \leq TH^o$.

(b) *If v_n^+ is isotone, then for all $(m,k) \in E$ holds*

$$v_n^o(m,k) \leq v_n^+(m,k) \quad \forall n \in \mathbb{N},$$

and consequently $TH^o \leq TH^+$.

(c) *If v_n^o is isotone, then for all $(m,k) \in E$ holds*

$$v_n^-(m,k) \leq v_n^o(m,k) \leq v_n^+(m,k) \quad \forall n \in \mathbb{N},$$

and consequently $TH^- \leq TH^o \leq TH^+$.

Proof. The equations can be proven by induction. The proof is presented in Appendix D.2 on page 349. □

Isotonicity of the finite time cumulative rewards is an intuitive property under any of the three ageing regimes: If we have either more inventory and at least the same number of customers in the system, or more customers in the system and at least the same stock size of the inventory the system should be able to produce more output.

The proof of this monotonicity does not seem to be direct, we have partial results only. The conditions $\mu = \gamma$ under ageing regime $k \mapsto (\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{m>0\}}$ in Proposition 4.3.13 and $\lambda \leq \gamma$ under ageing regime $k \mapsto \gamma \cdot k$ in Proposition 4.3.12 imply that ageing is in any case fast.

Proposition 4.3.12. *If in the production-inventory system with ageing regime $k \mapsto \gamma \cdot k$ we have $\lambda \leq \gamma$, then the finite time cumulative rewards $v_n^-(m,k)$ are isotone with respect to the natural order.*

We show by induction isotonicity in both directions and that the increase is bounded. For all $n \in \mathbb{N}$ holds

$$\begin{aligned}
v_n^-(m,k) - v_n^-(m,k-1) &\geq 0, && \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0,\\
v_n^-(m+1,k) - v_n^-(m,k) &\geq 0, && \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0,\\
v_n^-(m+1,k) - v_n^-(m,k) &\leq \alpha, && \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0,\\
v_n^-(m,k) - v_n^-(m,k-1) &\leq \alpha, && \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0.
\end{aligned}$$

The proof is presented in Appendix D.2 on page 354.

Proposition 4.3.13. *If in the production-inventory system with ageing regime* $k \mapsto (\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{m>0\}}$ *we have* $\mu = \gamma$*, then the finite time cumulative rewards* $v_n^o(m,k)$ *are isotone with respect to the natural order.*

Proof. We show by induction isotonicity in both directions, that the increase is bounded, and that v_n^o is concave in time. For all $n \in \mathbb{N}$ holds

$$\begin{aligned} v_n^o(m,k) - v_n^o(m,k-1) &\geq 0, && \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0,\\ v_n^o(m+1,k) - v_n^o(m,k) &\geq 0, && \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0,\\ v_n^o(m+1,k) - v_n^o(m,k) &\leq \alpha, && \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0,\\ v_n^o(m,k) - v_n^o(m,k-1) &\leq \alpha, && \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0,\\ v_n^o(m+1,k) - 2 \cdot v_n^o(m,k) + v_n^o(m-1,k) &\leq 0, && \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}. \end{aligned}$$

The proof is presented in Appendix D.2 on page 363. □

Location specific items

5. Basic production-inventory model with base stock policy

Parts of this chapter are taken from [OKD17].

5.1. Own contributions

We develop a Markov process model of a complex supply chain and derive its stationary distributions of the joint queueing-inventory process in explicit product form. This enables us to perform cost analysis and optimization of the system and to analyse the structure of the system in detail.

If we consider the production facilities (queues) at the locations as devices which deliver items from the inventory to incoming demand, needing non-negligible delivering time (as in the single-echelon inventory systems case described by Krishnamoorthy, Lakshmy, and Manikandan [KLM11]), our results extend their setting to a multi-dimensional system. On the other hand our work is an extension of the investigations of Rubio and Wein [RW96], Zazanis [Zaz94] and Reed and Zhang [RZ17] on inventory systems under base stock policy: In their models there is no production-to-order such that the time to satisfy customer demand is zero. Therefore, their model is a special case of our model when service time is set to zero.

Our system is an extension of [OKD16], which is presented in Chapter 2, and can be classified as a "multi-product system" because items are not exchangeable. Our network's behaviour, where the orders are dedicated to the sending locations, is more complicated and the supplier can be of a complex structure, e.g. a production network itself.

5.2. Description of the model

The supply chain of interest is depicted in Figure 5.2.1. We have a set $\overline{J} := \{1, 2, \dots, J\}$ of locations. Each location consists of a production system with an attached inventory. The inventories are replenished by a supplier network which consists of a set $\overline{M} := \{J+1, \dots, J+M\}$ of workstations and manufactures raw material for all locations, but distinguishes the replenishment orders from different locations. Each order for raw material is specified by a location $j \in \overline{J}$ and the resulting raw material is sent back exactly to the location which has placed the order.

Facilities in the supply chain. Each production system $j \in \overline{J}$ consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under FCFS regime. Customers arrive one by one at production system j according to a Poisson process with rate $\lambda_j > 0$ and require service. To satisfy a customer's demand the production system requires exactly one item of raw material, which is taken from the

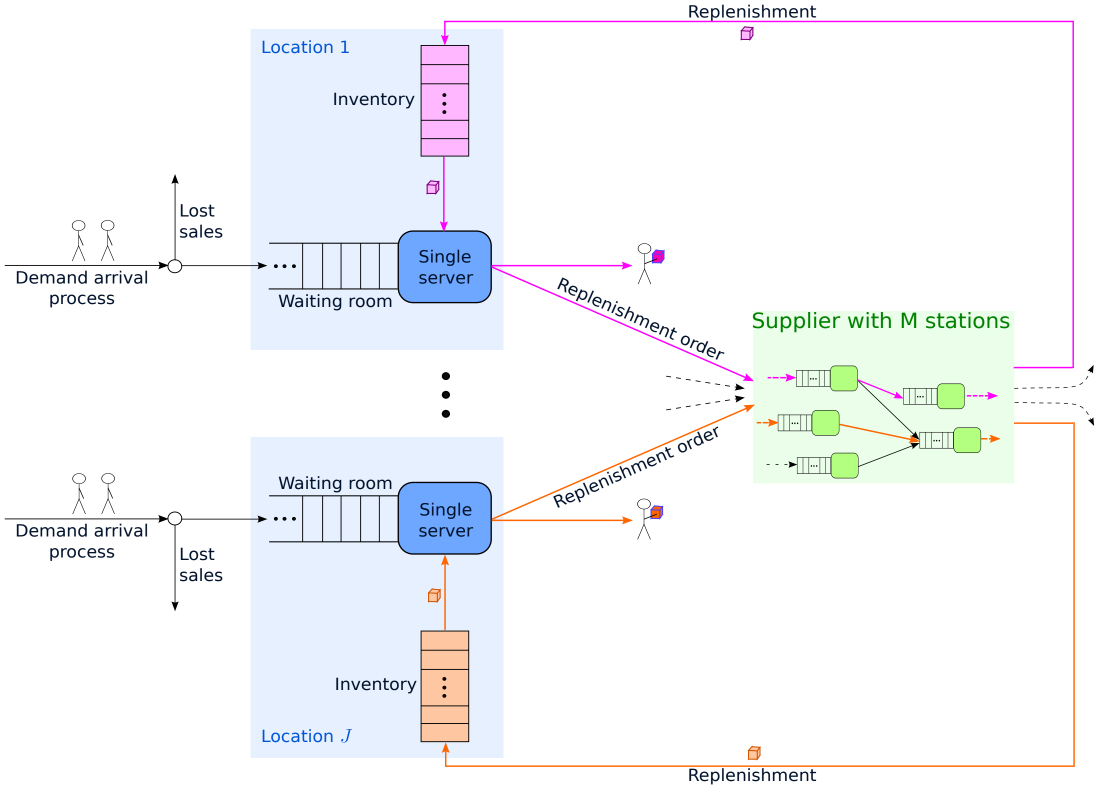


Figure 5.2.1.: Supply chain with base stock policy

associated local inventory. When a new customer arrives at a location while the previous customer's processing is not finished, this customer will wait. If inventory is depleted at location j, the customers who are already waiting in line will continue to wait, but newly arriving customers at this location will decide not to join the queue and are lost ("local lost sales").

The service requests at the locations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at location $j \in \overline{J}$ is provided with local queue-length-dependent intensity. If there are $n_j > 0$ customers present at location j either waiting or in service (if any) and if the inventory is not depleted, the service intensity is $\mu_j(n_j) > 0$. If the server is ready to serve a customer who is at the head of the line, and the inventory is not depleted, service immediately starts. Otherwise, the service starts at the instant when the next replenishment arrives at the local inventory.

The inventory at location j is controlled by prescribing a local base stock level $b_j \geq 1$, which is the maximal size of the inventory there, we denote $\mathbf{b} := \big(b_j : j \in \overline{J}\big)$.

Each workstation $m \in \overline{M}$ of the supplier network consists of a single server with infinite waiting room under FCFS regime. The service requests at the workstations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. Service at workstation $m \in \overline{M}$ is provided with local queue-length-dependent intensity. If there are $\ell > 0$ orders present, the service intensity is $\nu_m(\ell) > 0$.

Routing in the supply chain. A served customer departs from the system (with the consumed material) immediately after service and at the same time an order for one item of raw material is placed at the supplier network ("base stock policy").

To distinguish orders from different locations, each order is marked (tagged) by a "type" which for simplicity is the index of the location, where the order is triggered. We found that Kelly's deterministic routing scheme for "customers" in networks (cf. [Kel79, pp. 82ff.]) is a useful device to describe the interaction of inventories and supplier network. It should be emphasized that the cycling "customer" represents an order in the supplier network and a item of raw material in the inventories.
An order triggered by location j follows a type-j-dependent route for eventual replenishment, denoted by $r(j) = \Big(r(j,1), \ldots, r(j, S(j)-1), r(j, S(j))\Big)$. Here $r(j,\ell) \in \overline{M}$ for $\ell = 2, \ldots, S(j)$ is the identifier of the ℓ-th workstation on the path $r(j)$, and $S(j)$ is the number of stages of the route of type j. For completeness we fix $r(j,1) := j \in \overline{J}$, and prescribe that a type-j order departing from $r(j, S(j))$ enters as an item of raw material immediately the inventory at location $j = r(j,1)$ to restart its cycle.

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the supplier network and the local inventories are negligible.

The usual independence assumptions are assumed to hold as well.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the contents of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_m(t)$ we denote the sequence of orders at workstation $m \in \overline{M}$ of the supplier network at time $t \geq 0$.

We denote by K_m the set of possible states at node $m \in \overline{M}$ (local state space). The state $k_m := [t_{m1}, s_{m1}; \ldots; t_{m\#k_m}, s_{m\#k_m}] \in K_m$ indicates that there are $\#k_m$ orders at workstation $m \in \overline{M}$, on position $p \in \{1, \ldots, \#k_m\}$ resides an order of type $t_{mp} \in \overline{J}$, which is on stage $s_{mp} \in \{1, \ldots, S(t_{mp})\}$ of its route $r(t_{mp}) = \Big(r(t_{mp}, 1), \ldots, r(t_{mp}, S(t_{mp}))\Big)$. Here (t_{m1}, s_{m1}) is the order at the head of the line, which is in service and $(t_{m\#k_m}, s_{m\#k_m})$ is the order at the tail of the line.

Notational convention. To make reading easier, we use a unified notation for the states of the inventories at the locations and the states of the workstations in the supplier network. In doing this we identify items of raw material arriving at the inventory j with the order sent out to the supplier network when an item is consumed by a departing customer. Therefore, adopting the state description of the workstations for that of the inventories, the state of the inventory at location $j \in \overline{J}$ at time t is

$$Y_j(t) = k_j = \underbrace{[j,1; \ldots; j,1]}_{\#k_j \text{ items}},$$

since the route of type j starts in the inventory at location j (i.e. $t_{jp} = j$ for the types and $s_{jp} = 1$ for all $p \in \{1, \ldots, \#k_j\}$). A stage number $s_{jp} > 1$ indicates that the unit (as an order) is in the supplier network.

Summarizing, global states of the inventory-replenishment subsystem are

$$\mathbf{k} = \Big(\overbrace{k_1, \ldots, k_J}^{\substack{\text{inventories} \\ \text{at locations}}}, \overbrace{k_{J+1}, \ldots, k_{J+M}}^{\substack{\text{workstations} \\ \text{at supplier network}}}\Big) \in K \subseteq \prod_{j=1}^{J+M} K_j,$$

where K_j denotes the local state space at $j \in \overline{J} \cup \overline{M}$ and K denotes the feasible states composed of feasible local states.

For $\#k_j = 0$, $j \in \overline{J}$, we read

$$\left[t_{j1}, s_{j1}; \ldots; t_{j\#k_j}, s_{j\#k_j}\right] =: [0],$$

and for $\#k_m = 0$, $m \in \overline{M}$, we read

$$[t_{m1}, s_{m1}; \ldots; t_{m\#k_m}, s_{m\#k_m}] =: [0].$$

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), \ldots, X_J(t), Y_1(t), \ldots, Y_J(t), W_{J+1}(t), \ldots, W_{J+M}(t)) : t \geq 0)\,.$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \left\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J, \mathbf{k} \in K\right\}.$$

Discussion of the modelling assumptions. We have imposed simplifying assumptions on the production-inventory system to obtain explicit and simple-to-calculate performance metrics of the system, which give insights into its long-time and stationary behaviour. This enables a parametric and sensitivity analysis that is easy to perform.

First, the assumption of exponentially distributed inter-arrival and service times are standard in the literature. The locally state-dependent service rates are also common and give quite a bit of flexibility. The lead time of an order is composed of the waiting times plus the service times in the supplier network. They are therefore more complex than exponential or even constant lead times.

Second, we assume that the local base stock levels are positive (i.e. $b_j \geq 1$ at location j). If $b_j = 0$, all customers at location j would be lost, which is the same as excluding location j from the production-inventory system.

Third, the assumption of zero transportation times can be removed by inserting special (virtual) $M/G/\infty$ workstations into the network.

5.3. Limiting and stationary behaviour

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n}, \mathbf{k}) \in E$, where a typical state is (we will impose necessary restrictions if needed)

$$(\mathbf{n},\mathbf{k}) = \Big(\mathbf{n}, \underbrace{[1,1;\dots;1,1]}_{\#k_1}, \dots, \underbrace{[J,1;\dots;J,1]}_{\#k_J}, \big[t_{(J+1)1}, s_{(J+1)1}; \dots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\big], \dots \\ \dots, \big[t_{(J+M)1}, s_{(J+M)1}; \dots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\big]\Big). \tag{5.3.1}$$

- ARRIVAL OF A CUSTOMER AT LOCATION $i \in \overline{J}$,
  which happens only if the inventory at this location is not empty because of the lost sales rule:
  $$q\left((\mathbf{n},\mathbf{k});(\mathbf{n}+\mathbf{e}_i,\mathbf{k})\right) = \lambda_i \cdot 1_{\{\#k_i>0\}}, \qquad i \in \overline{J}.$$

- SERVICE COMPLETION OF A CUSTOMER AT LOCATION $i \in \overline{J}$,
  which happens only if there is at least one customer at location i and the inventory there is not empty,
  i.e. from location i (=station $r(i,1)$), where $\#k_i$ items are present, a customer departs and an item of raw material is removed from the associated local inventory, in addition a replenishment order is sent to workstation $r(i,2) \in \overline{M}$ of the supplier network, where $\#k_{r(i,2)}$ orders have already been present:

$$q\bigg(\Big(\mathbf{n}, \underbrace{[1,1;\dots;1,1]}_{\#k_1}, \dots, \underbrace{[i,1;\dots;i,1]}_{\#k_i>0}, \dots, \underbrace{[J,1;\dots;J,1]}_{\#k_J}, \big[t_{(J+1)1}, s_{(J+1)1}; \dots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\big], \dots \\ \dots, \underbrace{\Big[t_{r(i,2)1}, s_{r(i,2)1}; \dots; t_{r(i,2)\#k_{r(i,2)}}, s_{r(i,2)\#k_{r(i,2)}}\Big]}_{\#k_{r(i,2)}}, \dots \\ \dots, \big[t_{(J+M)1}, s_{(J+M)1}; \dots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\big]\Big); \\ \Big(\mathbf{n}-\mathbf{e}_i, \underbrace{[1,1;\dots;1,1]}_{\#k_1}, \dots, \underbrace{[i,1;\dots;i,1]}_{\#k_i-1}, \dots, \underbrace{[J,1;\dots;J,1]}_{\#k_J}, \big[t_{(J+1)1}, s_{(J+1)1}; \dots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\big], \dots \\ \dots, \underbrace{\Big[t_{r(i,2)1}, s_{r(i,2)1}; \dots; t_{r(i,2)\#k_{r(i,2)}}, s_{r(i,2)\#k_{r(i,2)}}; i, 2\Big]}_{\#k_{r(i,2)}+1}, \dots \\ \dots, \big[t_{(J+M)1}, s_{(J+M)1}; \dots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\big]\Big)\bigg) \tag{5.3.2}$$
$$= \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}, \qquad i \in \overline{J}.$$

Notational convention: For transition rates like the one above we will henceforth use an abbreviated notation. Using this abbreviation (5.3.2) reads

$$q\Bigg(\Big(\mathbf{n},\ldots,\underbrace{[i,1;\ldots;i,1]}_{\#k_i>0},\ldots,\underbrace{\Big[t_{r(i,2)1},s_{r(i,2)1};\ldots;t_{r(i,2)\#k_{r(i,2)}},s_{r(i,2)\#k_{r(i,2)}}\Big]}_{\#k_{r(i,2)}},\ldots\Big);$$
$$\Big(\mathbf{n}-\mathbf{e}_i,\ldots,\underbrace{[i,1;\ldots;i,1]}_{\#k_i-1},\ldots,\underbrace{\Big[t_{r(i,2)1},s_{r(i,2)1};\ldots;t_{r(i,2)\#k_{r(i,2)}},s_{r(i,2)\#k_{r(i,2)}};i,2\Big]}_{\#k_{r(i,2)}+1},\ldots\Big)\Bigg)$$
$$=\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}},\qquad i\in\overline{J}.$$

This means that we will explicitly write only those local states of $(\mathbf{n},\mathbf{k})$ and its successor state which are relevant for the described transition. Readers interested in the full expressions may consult [OKD14].

- Service completion of an order at workstation $m\in\overline{M}$,
which happens only if there is at least one order,
i.e. from workstation m, where $\#k_m$ orders are present, an order of type t_{m1} on stage s_{m1} of its route is removed and is sent to the next stage $s_{m1}+1$ of its route, i.e. either

- if $s_{m1}<S(t_{m1})$, it is sent to workstation $r(t_{m1},s_{m1}+1)\in\overline{M}$, where $\#k_{r(t_{m1},s_{m1}+1)}$ orders have already been present:

$$q\Bigg(\Big(\mathbf{n},\ldots,\underbrace{[t_{m1},s_{m1};\ldots;t_{m\#k_m},s_{m\#k_m}]}_{\#k_m>0},\ldots$$
$$\ldots,\underbrace{\Big[t_{r(t_{m1},s_{m1}+1)1},s_{r(t_{m1},s_{m1}+1)1};\ldots;t_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}},s_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}}\Big]}_{\#k_{r(t_{m1},s_{m1}+1)}},\ldots\Big);$$
$$\Big(\mathbf{n},\ldots,\underbrace{[t_{m2},s_{m2};\ldots;t_{m\#k_m},s_{m\#k_m}]}_{\#k_m-1},\ldots$$
$$\ldots,\underbrace{\Big[t_{r(t_{m1},s_{m1}+1)1},s_{r(t_{m1},s_{m1}+1)1};\ldots;t_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}},s_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}};t_{m1},s_{m1}+1\Big]}_{\#k_{r(t_{m1},s_{m1}+1)}+1},\ldots\Big)\Bigg)$$
$$=\nu(\#k_m)\cdot 1_{\{\#k_m>0\}}\cdot 1_{\{s_{m1}<S(t_{m1})\}},\qquad m\in\overline{M},$$

- or if $s_{m1} = S(t_{m1})$, it is sent to the inventory at location $t_{m_1} \in \overline{J}$, where $\#k_{t_{m1}}$ items of raw material have already been present:

$$q\Bigg(\Big(\mathbf{n}, \ldots, \underbrace{[t_{m1}, 1; \ldots; t_{m1}, 1]}_{\#k_{t_{m1}}}, \ldots, \underbrace{[t_{m1}, s_{m1}; \ldots; t_{m\#k_m}, s_{m\#k_m}]}_{\#k_m > 0}, \ldots\Big); \\ \Big(\mathbf{n}, \ldots, \underbrace{[t_{m1}, 1; \ldots; t_{m1}, 1; t_{m1}, 1]}_{\#k_{t_{m1}}+1}, \ldots, \underbrace{[t_{m2}, s_{m2}; \ldots; t_{m\#k_m}, s_{m\#k_m}]}_{\#k_m - 1}, \ldots\Big)\Bigg) \\ = \nu(\#k_m) \cdot 1_{\{\#k_m > 0\}} \cdot 1_{\{s_{m1} = S(t_{m1})\}}, \qquad m \in \overline{M}.$$

Furthermore, $q(z; \tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q(z; z) = -\sum_{\substack{\tilde{z} \in E, \\ z \neq \tilde{z}}} q(z; \tilde{z}) \qquad \forall z \in E.$$

Proposition 5.3.1. *The strictly positive measure* $\mathbf{x} := (x(\mathbf{n}, \mathbf{k}) : (\mathbf{n}, \mathbf{k}) \in E)$ *with*

$$x(\mathbf{n}, \mathbf{k}) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}(\mathbf{k}), \tag{5.3.3}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0, \ j \in \overline{J}, \tag{5.3.4}$$

$$\widetilde{\theta}(\mathbf{k}) = C_\theta^{-1} \prod_{j \in \overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m \in \overline{M}} \prod_{\ell=1}^{\#k_m} \frac{1}{\nu_m(\ell)}, \quad \mathbf{k} \in K, \tag{5.3.5}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z.

Proof. Before proving we recall notation for the inventory-replenishment subsystem: It will be convenient to use the elaborate although redundant notation for $\mathbf{k} \in K$

$$\mathbf{k} = \Big(\overbrace{k_1, \ldots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{k_{J+1}, \ldots, k_{J+M}}^{\substack{\text{workstations}\\ \text{at supplier network}}}\Big) \\ = \Bigg(\overbrace{\underbrace{[1, 1; \ldots; 1, 1]}_{\#k_1 \text{ items}}, \ldots, \underbrace{[J, 1; \ldots; J, 1]}_{\#k_J \text{ items}}}^{\substack{\text{inventories}\\ \text{at locations}}}, \\ \overbrace{\underbrace{[t_{(J+1)1}, s_{(J+1)1}; \ldots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}]}_{\#k_{J+1} \text{ orders}}, \ldots \\ \ldots, \underbrace{[t_{(J+M)1}, s_{(J+M)1}; \ldots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}]}_{\#k_{J+M} \text{ orders}}}^{\text{supplier network}}\Bigg).$$

The global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ of the queueing-inventory process Z are for $(\mathbf{n}, \mathbf{k}) \in E$ from (5.3.1) given by

flux out of the state $(\mathbf{n}, \mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if the inventory at this location is not empty (i.e. $\#k_i > 0$) because of the lost sales rule,
- a service completion of a customer at location $i \in \overline{J}$
  if there is at least one customer (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$),
- a completion of an order at workstation $\ell \in \overline{M}$ of the supplier network
  if there is at least one order at this workstation (i.e. $\#k_\ell > 0$):

$$x\Big(\mathbf{n}, \mathbf{k}\Big) \cdot \Big(\underline{\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{\#k_i > 0\}}} + \underline{\sum_{i \in \overline{J}} \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{\#k_i > 0\}}} + \underline{\sum_{\ell \in \overline{M}} \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell > 0\}}}\Big)$$

$=$ flux into the state $(\mathbf{n}, \mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one customer at location i (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$):

$$\sum_{i \in \overline{J}} x\Big(\mathbf{n} - \mathbf{e}_i, \mathbf{k}\Big) \cdot \lambda_i \cdot 1_{\{\#k_i > 0\}} \cdot 1_{\{n_i > 0\}}$$

- a service completion of a customer at location $t_{\ell \# k_\ell} = r(t_{\ell \# k_\ell}, 1) \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one order at workstation ℓ (i.e. $\#k_\ell > 0$) and the order at the tail of the queue at workstation ℓ is in stage 2 (i.e. $s_{\ell \# k_\ell} = 2$) of its route
  (i.e. a customer departs from location $t_{\ell \# k_\ell}$
  and an item is removed from the associated local inventory there,
  and an order is sent to workstation $r(t_{\ell \# k_\ell}, 2) = \ell$)
  (note that $s_{\ell \# k_\ell} = 2$ implies $\#k_{t_{\ell \# k_\ell}} < b_{t_{\ell \# k_\ell}}$ to hold):

$$+ \sum_{\ell \in \overline{M}} x\Big(\mathbf{n} + \mathbf{e}_{t_{\ell \# k_\ell}}, \ldots, \underbrace{[t_{\ell \# k_\ell}, 1; t_{\ell \# k_\ell}, 1; \ldots; t_{\ell \# k_\ell}, 1]}_{\#k_{t_{\ell \# k_\ell}} + 1}, \ldots$$

$$\ldots, \underbrace{[t_{\ell 1}, s_{\ell 1}; \ldots; t_{\ell(\#k_\ell - 1)}, s_{\ell(\#k_\ell - 1)}]}_{\#k_\ell - 1}, \ldots\Big)$$

$$\cdot \mu_{t_{\ell \# k_\ell}}(n_{t_{\ell \# k_\ell}} + 1) \cdot 1_{\{s_{\ell \# k_\ell} = 2\}} \cdot 1_{\{\#k_\ell > 0\}}$$

- a transition of an order of type $t_{\ell\#k_\ell}$ from workstation $r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)$
  to the next workstation of the supplier network
  if in state $(\mathbf{n},\mathbf{k})$ there is at least one order at workstation ℓ (i.e. $\#k_\ell > 0$) and the order in the tail of the queue at workstation ℓ is not in stage 2 (i.e. $s_{\ell\#k_\ell} > 2$) of its route
  (i.e. an order of type $t_{\ell\#k_\ell}$ is removed from workstation $r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)$
  and is sent to workstation $\ell = r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell})$):

$$+\sum_{\ell\in\overline{M}} x\Big(\mathbf{n},\dots,\underbrace{\Big[t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1; t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1}, s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\dots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}, s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\dots$$
$$\dots,\underbrace{\big[t_{\ell 1}, s_{\ell 1};\dots;t_{\ell(\#k_\ell-1)}, s_{\ell(\#k_\ell-1)}\big]}_{\#k_{t_{\ell\#k_\ell}}-1},\dots\Big)$$
$$\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}$$

- a replenishment of the inventory at location $i \in \overline{J}$
  if in state $(\mathbf{n},\mathbf{k})$ there is at least one item of raw material at location i (i.e. $\#k_i > 0$)
  (i.e. an order of type i is removed from workstation $r(i, S(i))$
  and is sent to the inventory at location i):

$$+\sum_{i\in\overline{J}} x\Big(\mathbf{n},\dots,\underbrace{[i,1;\dots;i,1]}_{\#k_i-1},\dots$$
$$\dots,\underbrace{\Big[i,S(i);t_{r(i,S(i))1}, s_{r(i,S(i))1};\dots;t_{r(i,S(i))\#k_{r(i,S(i))}}, s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\dots\Big)$$
$$\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i>0\}}.$$

It has to be shown that the stationary measure (5.3.3) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

Substitution of (5.3.3) and (5.3.4) into the global balance equations directly leads to

$$
\begin{aligned}
&\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\\
&\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}+\sum_{\ell\in\overline{M}}\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\Big)\\
=&\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}\setminus\{i\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_i(n_i-1)\cdot\widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}\\
&+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}\setminus\{t_{\ell\#k_\ell}\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_{t_{\ell\#k_\ell}}(n_{t_{\ell\#k_\ell}}+1)\\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[t_{\ell\#k_\ell},1;t_{\ell\#k_\ell},1;\ldots;t_{\ell\#k_\ell},1]}_{\#k_{t_{\ell\#k_\ell}}+1},\ldots,\underbrace{[t_{\ell1},s_{\ell1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\ldots\Big)\\
&\quad\cdot\mu_{t_{\ell\#k_\ell}}(n_{t_{\ell\#k_\ell}}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}=2\}}\\
&+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\underbrace{\Big[t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\ldots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\ldots\\
&\qquad\ldots,\underbrace{[t_{\ell1},s_{\ell1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_{t_{\ell\#k_\ell}}-1},\ldots\Big)\\
&\quad\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[i,1;\ldots;i,1]}_{\#k_i-1},\ldots\\
&\qquad\ldots,\underbrace{\Big[i,S(i);t_{r(i,S(i))1},s_{r(i,S(i))1};\ldots;t_{r(i,S(i))\#k_{r(i,S(i))}},s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\ldots\Big)\\
&\quad\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i>0\}}.
\end{aligned}
$$

By substitution of (5.3.4) we obtain

$$
\begin{aligned}
&\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k}) \\
&\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}+\sum_{\ell\in\overline{M}}\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\Big) \\
=&\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}} \\
&+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right) \\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[t_{\ell\#k_\ell},1,t_{\ell\#k_\ell},1;\ldots;t_{\ell\#k_\ell},1]}_{\#k_{t_{\ell\#k_\ell}}+1},\ldots,\underbrace{[t_{\ell1},s_{\ell1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\ldots\Big) \\
&\quad\cdot\lambda_{t_{\ell\#k_\ell}}\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}=2\}} \\
&+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right) \\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\Big[t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\ldots \\
&\qquad\underbrace{\ldots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\ldots \\
&\qquad\ldots,\underbrace{[t_{\ell1},s_{\ell1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_{t_{\ell\#k_\ell}}-1},\ldots\Big) \\
&\quad\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}} \\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right) \\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[i,1;\ldots;i,1]}_{\#k_i-1},\ldots \\
&\qquad\ldots,\underbrace{\Big[i,S(i);t_{r(i,S(i))1},s_{r(i,S(i))1};\ldots;t_{r(i,S(i))\#k_{r(i,S(i))}},s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\ldots\Big) \\
&\quad\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i>0\}}.
\end{aligned}
$$

Cancelling $\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$ and the sums with the terms $\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}$ on both sides of the equation leads to

$$\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{\ell\in\overline{M}}\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\Big) \tag{5.3.6}$$

$$=\sum_{\ell\in\overline{M}}\widetilde{\theta}\Big(\ldots,\underbrace{[t_{\ell\#k_\ell},1,t_{\ell\#k_\ell},1;\ldots;t_{\ell\#k_\ell},1]}_{\#k_{t_{\ell\#k_\ell}}+1},\ldots,\underbrace{[t_{\ell 1},s_{\ell 1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\ldots\Big)$$

$$\cdot\lambda_{t_{\ell\#k_\ell}}\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}=2\}}$$

$$+\sum_{\ell\in\overline{M}}\widetilde{\theta}\Big(\ldots,\Big[t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\ldots$$

$$\underbrace{\ldots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\ldots$$

$$\ldots,\underbrace{[t_{\ell 1},s_{\ell 1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_{t_{\ell\#k_\ell}}-1},\ldots\Big)$$

$$\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}$$

$$+\sum_{i\in\overline{J}}\widetilde{\theta}\Big(\ldots,\underbrace{[i,1;\ldots;i,1]}_{\#k_i-1},\ldots$$

$$\ldots,\underbrace{\Big[i,S(i);t_{r(i,S(i))1},s_{r(i,S(i))1};\ldots;t_{r(i,S(i))\#k_{r(i,S(i))}},s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\ldots\Big)$$

$$\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i>0\}}.$$

Now, substitution of (5.3.5) leads to

$$\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{\ell\in\overline{M}}\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\Big)$$

$$=\sum_{\ell\in\overline{M}}\widetilde{\theta}(\mathbf{k})\cdot\left(\frac{1}{\lambda_{t_{\ell\#k_\ell}}}\right)\cdot\nu_\ell(\#k_\ell)\cdot\lambda_{t_{\ell\#k_\ell}}\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}=2\}}$$

$$+\sum_{\ell\in\overline{M}}\widetilde{\theta}(\mathbf{k})\cdot\left(\frac{1}{\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)}\right)\cdot\nu_\ell(\#k_\ell)$$

$$\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}$$

$$+\sum_{i\in\overline{J}}\widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot\left(\frac{1}{\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)}\right)\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i>0\}}.$$

The right-hand side of the last equation is

$$\sum_{\ell\in\overline{M}}\widetilde{\theta}(\mathbf{k})\cdot\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}=2\}}+\sum_{\ell\in\overline{M}}\widetilde{\theta}(\mathbf{k})\cdot\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}$$

$$+\sum_{i\in\overline{J}}\widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot 1_{\{\#k_i>0\}},$$

which is obviously the left-hand side.

Inspection of the system (5.3.6) reveals that it is a "generator equation", i.e. the global balance equation $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$ for a suitably defined ergodic Markov process on state space K with "reduced generator" $\mathbf{Q}_{red}$. Because the Markov process generated by $\mathbf{Q}_{red}$ is irreducible the solution of (5.3.6) is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Remark 5.3.2. The system (5.3.6) of equations resembles the global balance equations of a closed Kelly network with $J+M$ nodes and $\sum_{j\in\overline{J}} b_j$ customers and exponentially distributed service times with rate λ_j at node $j \in \{1,\ldots,J\}$ and with queue-length-dependent rate $\nu_m(\cdot)$ at node $m \in \{J+1,\ldots,J+M\}$ and with deterministic, type-dependent routing.

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 5.3.3. For the queueing-inventory system Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E), \quad \pi(\mathbf{n},\mathbf{k}) := \lim_{t\to\infty} P(Z(t) = (\mathbf{n},\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := \left(\xi(\mathbf{n}) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi(\mathbf{n}) := \lim_{t\to\infty} P((X_1(t),\ldots,X_J(t)) = \mathbf{n}),$$

$$\theta := (\theta(\mathbf{k}) : \mathbf{k} \in K), \quad \theta(\mathbf{k}) := \lim_{t\to\infty} P((Y_1(t),\ldots,Y_J(t),W_{J+1}(t),\ldots,W_{J+M}(t)) = \mathbf{k}).$$

Theorem 5.3.4. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n},\mathbf{k}) = \xi(\mathbf{n}) \cdot \theta(\mathbf{k}), \tag{5.3.7}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{5.3.8}$$

$$\theta(\mathbf{k}) = C_\theta^{-1} \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m\in\overline{M}} \prod_{\ell=1}^{\#k_m} \frac{1}{\nu_m(\ell)}, \quad \mathbf{k} \in K, \tag{5.3.9}$$

and normalisation constants

$$C_j = \sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} \quad \text{and} \quad C_\theta = \sum_{\mathbf{k}\in K} \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m\in\overline{M}} \prod_{\ell=1}^{\#k_m} \frac{1}{\nu_m(\ell)}.$$

Proof. Z is ergodic, if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ from Proposition 5.3.1 can be normalised (i.e. $\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n}, \mathbf{k}) < \infty$). Because of Proposition 5.3.1 it holds

$$\begin{aligned}\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n}, \mathbf{k}) &= \sum_{\mathbf{n}\in\mathbb{N}_0} \widetilde{\xi}(\mathbf{n}) \cdot \sum_{\mathbf{k}\in K} \widetilde{\theta}(\mathbf{k}) \\ &= \left(\prod_{j\in\overline{J}} \sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} \right) \cdot \left(\sum_{\mathbf{k}\in K} \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m\in\overline{M}} \prod_{\ell=1}^{\#k_m} \frac{1}{\nu_m(\ell)} \right).\end{aligned}$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 5.3.1 can be normalised if and only if $\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty$ for all $j \in \overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n}, \mathbf{k}) = \frac{x(\mathbf{n}, \mathbf{k})}{\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n}, \mathbf{k})},$$

where $x(\mathbf{n}, \mathbf{k})$ is given in Proposition 5.3.1. □

Remark 5.3.5. The expression (5.3.7) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

The explicit formula (5.3.9) for θ shows that in fact there exists a three-term product structure, and that moreover the equilibrium for the integrated model is stratified. In the upper stratum, we have two independent vectors for production and inventory-replenishment, the latter splits into two products, a factor for the subsystem comprising the inventories and a factor for the replenishment subsystem.

In the lower stratum each of the three factors of the upper stratum is decomposed completely in "single-component" factors concerning the production servers, the inventories, and the replenishment servers. It should be noted that the factors for the inventories and the replenishment servers do not indicate internal independence, but they are of product form as the celebrated conditionally independent coordinates in the equilibrium of Gordon-Newell networks (see Theorem A.2.6). Remark 5.3.2 may explain to a certain extend the conditional independence inside of θ.

Representation (5.3.8) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(n) = \prod_{j\in\overline{J}} \xi_j(n_j) = \prod_{j\in\overline{J}} C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. In our production network, however, servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (5.3.7) has been unexpected to us.

Our production-inventory-replenishment system can be considered as a "Jackson network in a random environment" in [KDO16, Section 4]. We can interpret the inventory-replenishment subsystem, which contributes via θ to Theorem 5.3.4, as a "random environment" for the production network of nodes $\overline{J}$, which is a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1] we conclude from the hindsight that decoupling of the queueing process $(X_1, \dots, X_J)$ and the process $(Y_1, \dots, Y_J, W_{J+1}, \dots, W_{J+M})$, i.e. the formula (5.3.7), is a consequence of that Theorem 4.1. It should be noted that the three-term structure of the upper stratum of the product form steady state π in Theorem 5.3.4 can not be obtained from the general theory.

Our direct proof of Theorem 5.3.4 is much shorter than embedding the present model into the general framework of [KDO16].

The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

5.4. Cost analysis

We consider the following cost structure for inventory, production, and replenishment. The total costs at location $j \in \overline{J}$ consist of shortage costs $c_{ls,j}$ for each lost customer, waiting costs $c_{w,j}$ per unit of time for each customer in the system (waiting or in service), capacity costs $c_{s,j}$ per unit of time for providing inventory storage space (e.g. rent, insurance), holding costs $c_{h,j}$ per unit of time for each unit that is kept on inventory. The unit holding costs per item at workstation $m \in \overline{M}$ of the supplier network are $c_{h,m}$. All $c_{ls,j}$, $c_{w,j}$, $c_{s,j}$, $c_{h,j}$, $c_{h,m}$ are positive.

Therefore, the global cost function $f_{\mathbf{b}} : \mathbb{N}_0^J \times K \longrightarrow \mathbb{R}_0^+$ per unit of time is

$$f_{\mathbf{b}}(\mathbf{n}, \mathbf{k}) = \left(\sum_{j \in \bar{J}} f_{b_j}(n_j, k_j) + \sum_{m \in \overline{M}} f_m(k_m) \right) \tag{5.4.1}$$

with cost functions $f_{b_j} : \mathbb{N}_0 \times K_j \longrightarrow \mathbb{R}_0^+$ per unit of time at location j with base stock level b_j

$$f_{b_j}(n_j, k_j) = c_{w,j} \cdot n_j + c_{s,j} \cdot b_j + c_{h,j} \cdot \#k_j + c_{ls,j} \cdot \lambda_j \cdot 1_{\{\#k_j = 0\}}$$

and $f_m : K_m \longrightarrow \mathbb{R}_0^+$ at workstation m of the supplier network per unit of time

$$f_m(k_m) = c_{h,m} \cdot \#k_m.$$

We are interested in the long time average costs of the system as a function of the base stock levels $\mathbf{b} = (b_1, \dots, b_J)$, which are considered as the main decision variables.

Lemma 5.4.1. *Optimal solutions for minimizing the asymptotic average costs with* (5.4.1) *are the elements in the set*

$$\arg\min\left(\overline{g}(\mathbf{b})\right)$$

with

$$\overline{g}(\mathbf{b}) := \sum_{j \in \bar{J}} c_{s,j} \cdot b_j + \sum_{\mathbf{k} \in K} \left(\sum_{j \in \bar{J}} c_{ls,j} \cdot \lambda_j \cdot 1_{\{\#k_j = 0\}} + \sum_{j \in \overline{J} \cup \overline{M}} c_{h,j} \cdot \#k_j \right) \cdot \theta(\mathbf{k}).$$

Proof. The structure of the proof is analogue to the proof of Lemma 2.5.1 whereby the costs for the workstations in the supplier network have to be considered. A detailed proof is presented in [OKD14, Lemma 5, pp. 16f.]. □

6. Supplier with symmetric server

In this chapter, we look at the aggregation of the supplier network. We can substitute the complex supplier network of Chapter 5 by only one node — a supplier who consists of a symmetric server.

6.1. Related literature and own contributions

The definition of the symmetric server follows Kelly [Kel79, Chapter 3] and is a well-known service discipline in network theory. Kelly's symmetric server is a generalization and unification of the nodes that are used to built the BCMP networks, which allow for non-exponential service times (cf. [Dad01b, Remark 9.5, p. 349]):

- Processor sharing:
  The capacity of the server is equally shared between the customers.
- Last-come, first-served preemptive resume (LCFS-PR):
  A newly arriving customer interrupts immediately an ongoing service. The preempted customer has to wait until the service of the newly arrived customer and his descendants is finished. Then the service of the preempted customer is resumed.
- Infinite server:
  There are infinitely many servers, so newly arriving customers do not need to wait for a server.

The symmetric server enables to deal with non-exponential type-dependent service time distribution for different order types.

Our reduction of the supplier network in the supply chain is analogous to Norton's theorem proved by Chandy et al. [CHW75]. They construct an "equivalent" network in which all the "uninteresting" queues outside of a predetermined subnetwork of special interest are replaced by one composite queue with a FCFS (or processor sharing) discipline and an appropriate service rate such that the behaviour of the subsystem of special interest in the equivalent network is identical with those in the original network. Norton's theorem holds for certain classes of queueing networks that satisfy local balance.

Towsley ([Tow75, Section 4.5, pp. 67ff.], [Tow80, Section 5, pp. 331ff.]) presents it in an extended form for locally balanced networks with state-dependent routing. Further extensions are summarised by Huisman and Boucherie [HB11, pp. 315ff.].

The critical view of Balsamo and Isazeolla [BI83] on the exact aggregation shows that the aggregation does not introduce computational savings in parametric analysis.

Our main contributions are the following:
We develop a Markov process model of a complex supply chain and derive its stationary distribution of the joint queueing-inventory process in explicit product form. A cost analysis and an eventual optimization can be performed as for the basic model in Section 5.4 on page 137. The symmetric server is a versatile model for the supplier network. To the best of our knowledge, symmetric servers have not been considered so far in the context of complex supply chains.

6.2. Description of the general model

The supply chain of interest is depicted in Figure 6.2.1. We have a set of locations $\overline{J} := \{1, 2, \ldots, J\}$. Each of the locations consists of a production system with an attached inventory. The inventories are replenished by a single central supplier, which is referred to as workstation $J+1$ and manufactures raw material for all locations, but distinguishes between the replenishment orders from different locations. Each order of raw material is specified by a location $j \in \overline{J}$ and the resulting raw material is sent to the location which has placed the order.

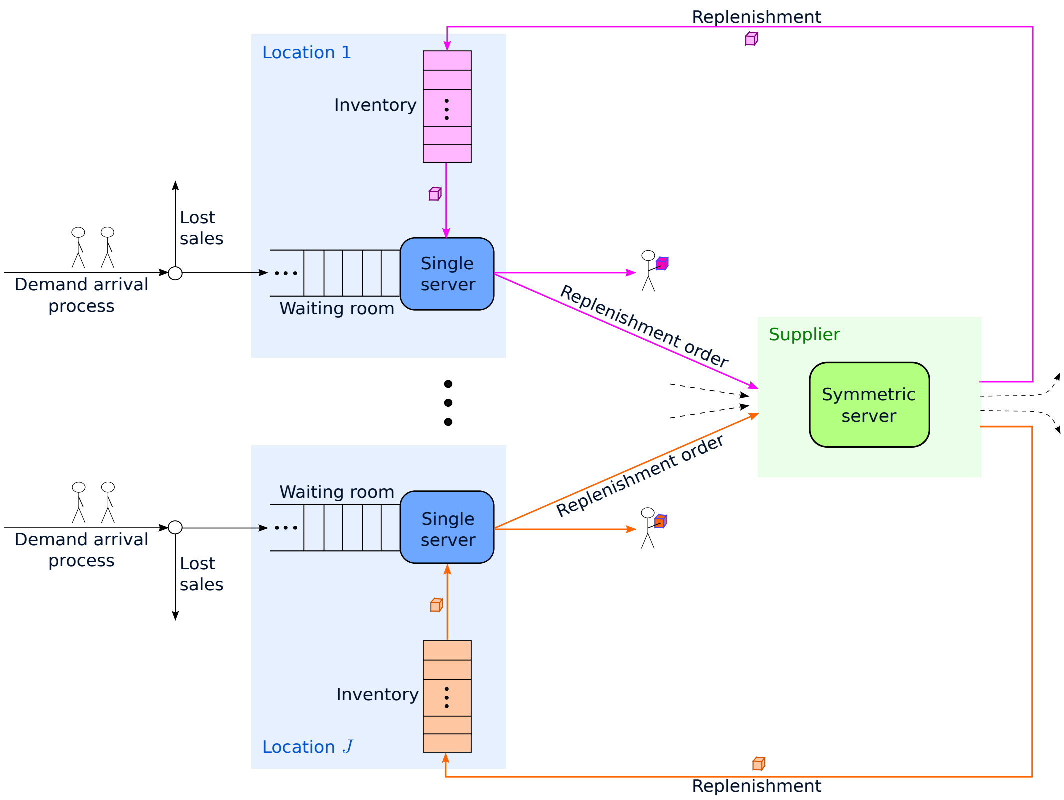


Figure 6.2.1.: Supply chain with base stock policy

Facilities in the supply chain. Each production system $j \in \overline{J}$ consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under a first-come, first-served (FCFS) regime. Customers arrive one by one at production system j according to a Poisson process with rate $\lambda_j > 0$ and require service.

To satisfy a customer's demand the production system needs exactly one item of raw material, which is taken from the associated local inventory. When a new customer arrives at a location while the previous customer's order is not finished, this customer will wait. If the inventory is depleted at location j, the customers who are already waiting in line will wait, but new arriving customers at this location will decide not to join the queue and are lost ("local lost sales").

The service requests at the locations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at location $j \in \overline{J}$ is provided with local queue-length-dependent intensity. If there are $n_j > 0$ customers present at location j either waiting or in service (if any) and if the inventory is not depleted, the service intensity is $\mu_j(n_j) > 0$. If the server is ready to serve a customer who is at the head of the line, and the inventory is not depleted, the service immediately starts. Otherwise, the service starts at the instant of time when the next replenishment arrives at the local inventory.

The inventory at location $j \in \overline{J}$ is controlled by prescribing a local base stock level $b_j \geq 1$, which is the maximal size of the inventory there, we denote $\mathbf{b} := \big(b_j : j \in \overline{J}\big)$.

The central supplier (which is referred to as workstation $J+1$) consists of a symmetric server. The systems under investigation differ with respect to the service time distribution of the central supplier in the following way:

- phase-type distribution (Section 6.3),
- exponential distribution (Section 6.4), whereby it is a special case of the phase-type distribution.

Routing in the supply chain. A served customer departs from the system immediately, and the associated consumed raw material is removed from the inventory, and an order of one unit is placed at the central supplier at this time instant ("base stock policy").

To distinguish orders from different locations, we mark each order by a "type" which for simplicity is the index of the location where the order is triggered.

The symmetric service discipline to be considered follows Kelly [Kel79, Chapter 3] and is defined as follows: There is a queue with $\sum_{j\in\overline{J}} b_j$ positions, where orders may reside. The positions are numbered $1, 2, \ldots$, and if there are $\#k_{J+1} > 0$ orders at the central supplier (either waiting or in service), they occupy positions $1, 2, \ldots, \#k_{J+1}$.

If there are $\#k_{J+1}$ orders present, then the central supplier offers a service capacity $\phi(\#k_{J+1}) > 0$, $\phi(0) = 0$. This service capacity is allocated to orders at the central supplier according to some function $c(\cdot, \#k_{J+1})$: The order on position p yields a portion $c(p, \#k_{J+1}) \in [0, 1]$ of the offered service capacity, $\sum_{p=1}^{\#k_{J+1}} c(p, \#k_{J+1}) = 1$.

If there are $\#k_{J+1}$ orders present and an arrival of a new order at the central supplier occurs, the new order is placed on some position $p \in \{1, 2, \ldots, \#k_{J+1}+1\}$ with probability $c(p, \#k_{J+1} + 1)$. Orders previously on positions $p, p + 1, \ldots, \#k_{J+1}$ are shifted to one step up into positions $p + 1, p + 2, \ldots, \#k_{J+1} + 1$.

If there are $\#k_{J+1} > 0$ orders present and the service time of the order on position $p \in \{1, \ldots, \#k_{J+1}\}$ expires, this order immediately departs from the central supplier and is sent to location $t_{(J+1)p}$, which has triggered the order. $t_{(J+1)p} \in \overline{J}$ is the type of the order on position p. Orders previously on positions $p+1, p+2, \dots, \#k_{J+1}$ are shifted one step down into positions $p, p+1, \dots, \#k_{J+1} - 1$.

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the central supplier and the local inventories are negligible.

The usual independence assumptions are assumed to hold as well.

6.3. Phase-type distributed service time

In this section, we study the queueing-inventory system as described in Section 6.2, where an order triggered by location $j \in \overline{J}$ requests for an amount of service time which is phase-type distributed. The mean service time request of type j is ν_j^{-1}.

We consider service time distributions of the following phase-type which are sufficiently versatile to approximate any distribution on $\mathbb{R}_0^+$ arbitrary close. The next definition is based on [Dad01b, Definition 9.2 , pp. 347f.].

Definition 6.3.1. For $h \in \mathbb{N}$ and $\beta_j > 0$ let

$$\Gamma_{\beta_j,h}(s) = 1 - e^{-\beta_j s} \sum_{i=0}^{h-1} \frac{(\beta_j s)^i}{i!}, \qquad s \geq 0,$$

denote the cumulative distribution function of the Γ-distribution with parameters β_j and h. The parameter h is a positive integer and serves as a phase-parameter for the number of independent exponential phases, each with mean β_j^{-1}, the sum of which constitutes a random variable with distribution $\Gamma_{\beta_j,h}$. ($\Gamma_{\beta_j,h}$ is called a h-stage Erlang distribution with scale parameter β_j.)

We consider the following class of distributions on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$, which is dense with respect to the topology of weak convergence of probability measures in the set of all distributions on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$ [Sch73, Section I.6]. For $\beta_j \in (0, \infty)$, $H_j \in \mathbb{N}$ and probability $b_j(\cdot)$ on $\{1, \dots, H_j\}$ with $b_j(H_j) > 0$ let the cumulative distribution function

$$B_j(s) = \sum_{\ell=1}^{H_j} b_j(\ell) \cdot \Gamma_{\beta_j,\ell}(s), \qquad s \geq 0,$$

denote a phase-type distribution function. With varying β_j, H_j and $b_j(\cdot)$ we can approximate any distribution on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$ sufficiently close.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the contents of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_{J+1}(t)$ we denote the sequence of orders at the central supplier at time $t \geq 0$.

We denote by K_{J+1} the set of possible states at the central supplier (local state space). The state

$$k_{J+1} := \underbrace{\left[t_{(J+1)1}, h_{(J+1)1}; \ldots; t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}\right]}_{\#k_{J+1} \text{ orders}} \in K_{J+1}$$

indicates that there are $\#k_{J+1}$ orders at the supplier and on position $p \in \{1, \ldots, \#k_{J+1}\}$ resides an order of type $t_{(J+1)p} \in \overline{J}$ requesting for exactly $h_{(J+1)p} \in \left\{1, \ldots, H_{t_{(J+1)p}}\right\}$ further independent exponential phases of service, each with mean $\beta^{-1}_{t_{(J+1)p}}$. More precisely, $t_{(J+1)1}$ is the order at the head of the line, that is in service and $t_{(J+1)\#k_{J+1}}$ is the order at the tail of the line.

Notational convention. To enhance readability, we use a unified notation for the states of the inventories at the locations and the state of the workstation at the central supplier. Therefore, if at time $t \geq 0$ the inventory size of $j \in \overline{J}$ is $Y_j(t) = \#k_j$, then the blown up state of the inventory at location $j \in \overline{J}$ is

$$k_j = \underbrace{[j; \ldots; j]}_{\#k_j \text{ items}} \in K_j.$$

The global states of the inventory-replenishment subsystem are then

$$\mathbf{k} = \Big(\overbrace{k_1, \ldots, k_J}^{\substack{\text{inventories} \\ \text{at locations}}}, \overbrace{k_{J+1}}^{\substack{\text{central} \\ \text{supplier}}}\Big) \in K \subseteq \prod_{j=1}^{J+1} K_j,$$

where K_j denotes the local state space at j, K denotes the feasible states composed of feasible local states.

For $\#k_j = 0$, $j \in \overline{J}$, we read

$$[j; \ldots; j] =: [0]$$

and for $\#k_{J+1} = 0$ we read

$$\left[t_{(J+1)1}, h_{(J+1)1}; \ldots; h_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}\right] =: [0].$$

We define the joint queueing-inventory process of this system by

$$Z = \left((X_1(t), \ldots, X_J(t), Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) \,:\, t \geq 0\right).$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \left\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J, \mathbf{k} \in K\right\}.$$

6.3.1. Limiting and stationary behaviour

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n},\mathbf{k}) \in E$, where a typical state is

$$(\mathbf{n},\mathbf{k}) = \Bigg(\mathbf{n}, \overbrace{\underbrace{[1;\dots;1]}_{\#k_1 \text{ items}}, \dots, \underbrace{[J;\dots;J]}_{\#k_J \text{ items}}}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{\underbrace{[t_{(J+1)1}, h_{(J+1)1}; \dots; h_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1} \text{ orders}}}^{\text{central supplier}}\Bigg) \tag{6.3.1}$$

and we will impose necessary restrictions if needed:

- Arrival of a customer at location $i \in \overline{J}$,
  which happens only if the inventory at this location is not empty because of the lost sales rule:
  $$q\left((\mathbf{n},\mathbf{k}); (\mathbf{n}+\mathbf{e}_i, \mathbf{k})\right) = \lambda_i \cdot 1_{\{\#k_i>0\}}, \qquad i \in \overline{J}.$$

- Service completion of a customer at location $i \in \overline{J}$,
  which happens only if there is at least one customer at location i and the inventory there is not empty,
  i.e. from location i, where $\#k_i > 0$ items are present, a customer departs and an item of raw material is removed from the associated local inventory,
  in addition a replenishment order is sent to the central supplier, where $\#k_{J+1}$ orders have been present, more precisely the order moves into position ℓ with probability $c(\ell, \#k_{J+1}+1)$ and orders previously in positions $\ell, \ell+1, \dots, \#k_{J+1}$ move to positions $\ell+1, \ell+2, \dots, \#k_{J+1}+1$ and the order has to obtain $\tilde{h}$ phases of service at the central supplier with probability $b_i(\tilde{h})$:

$$\begin{aligned}
q\Bigg(&\Big(\mathbf{n}, \underbrace{[1;\dots;1]}_{\#k_1}, \dots, \underbrace{[i;\dots;i]}_{\#k_i>0}, \dots, \underbrace{[J;\dots;J]}_{\#k_J},\\
&\qquad \underbrace{[t_{(J+1)1}, h_{(J+1)1}; \dots; t_{(J+1)\ell}, h_{(J+1)\ell}; \dots; t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}}\Big);\\
&\Big(\mathbf{n}-\mathbf{e}_i, \underbrace{[1;\dots;1]}_{\#k_1}, \dots, \underbrace{[i;\dots;i]}_{\#k_i-1}, \dots, \underbrace{[J;\dots;J]}_{\#k_J},\\
&\qquad \underbrace{[t_{(J+1)1}, h_{(J+1)1}; \dots; t_{(J+1)\ell-1}, h_{(J+1)\ell-1}; i, \tilde{h}; t_{(J+1)\ell}, h_{(J+1)\ell}; \dots; t_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}+1}\Big)\Bigg)\\
&= \mu_i(n_i) \cdot c(\ell, \#k_{J+1}+1) \cdot b_i(\tilde{h}) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}, \qquad i \in \overline{J},\ \ell \in \{1,\dots,\#k_{J+1}\}.
\end{aligned}$$

- Phase completion of an order on position ℓ at the central supplier, which happens only if there is at least one order, i.e. either

- if $h_{(J+1)\ell} = 1$, i.e. the order on position ℓ of type $t_{(J+1)\ell}$ is in its last phase of service,
i.e. from the central supplier, where $\#k_{J+1}$ orders are present, this order is removed and an item of raw material is sent to the inventory at location $t_{(J+1)\ell} \in \overline{J}$, where $\#k_{t_{(J+1)\ell}}$ items have already been present, in addition orders previously on positions $\ell, \ell+1, \ldots, \#k_{J+1}$ move to positions $\ell-1, \ell+1, \ldots, \#k_{J+1}-1$:

$$q\Bigg(\Big(\mathbf{n}, \underbrace{[1;\ldots;1]}_{\#k_1}, \ldots, \underbrace{[t_{(J+1)\ell};\ldots;t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}}, \ldots, \underbrace{[J;\ldots;J]}_{\#k_J}, \underbrace{[t_{(J+1)1}, h_{(J+1)1};\ldots;t_{(J+1)\ell-1}, h_{(J+1)\ell-1}; t_{(J+1)\ell}, 1; t_{(J+1)\ell+1}, h_{(J+1)\ell+1};\ldots;t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}>0}\Big);$$
$$\Big(\mathbf{n}, \underbrace{[1;\ldots;1]}_{\#k_1}, \ldots, \underbrace{[t_{(J+1)\ell};\ldots;t_{(J+1)\ell};t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}+1}, \ldots, \underbrace{[J;\ldots;J]}_{\#k_J}, \underbrace{[t_{(J+1)1}, h_{(J+1)1};\ldots;t_{(J+1)\ell-1}, h_{(J+1)\ell-1}; t_{(J+1)\ell+1}, h_{(J+1)\ell+1};\ldots;t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}-1}\Big)\Bigg)$$
$$= \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1}) \cdot c(\ell, \#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}}, \qquad \ell \in \{1, \ldots, \#k_{J+1}\},$$

- or if $h_{(J+1)\ell} > 1$, i.e. the service of the order on position ℓ of type $t_{(J+1)\ell}$ is not in its last phase of service, therefore the phase of this order is shifted one step down:

$$q\Bigg(\Big(\mathbf{n}, \underbrace{[1;\ldots;1]}_{\#k_1}, \ldots, \underbrace{[J;\ldots;J]}_{\#k_J}, \underbrace{[t_{(J+1)1}, h_{(J+1)1};\ldots;t_{(J+1)\ell-1}, h_{(J+1)\ell-1}; t_{(J+1)\ell}, h_{(J+1)\ell}; t_{(J+1)\ell+1}, h_{(J+1)\ell+1};\ldots;t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}>0}\Big);$$
$$\Big(\mathbf{n}, \underbrace{[1;\ldots;1]}_{\#k_1}, \ldots, \underbrace{[J;\ldots;J]}_{\#k_J}, \underbrace{[t_{(J+1)1}, h_{(J+1)1};\ldots;t_{(J+1)\ell-1}, h_{(J+1)\ell-1}; t_{(J+1)\ell}, h_{(J+1)\ell}-1; t_{(J+1)\ell+1}, h_{(J+1)\ell+1};\ldots;t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}>0}\Big)\Bigg)$$
$$= \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1}) \cdot c(\ell, \#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\{h_{(J+1)\ell}>1\}}, \ \ell \in \{1, \ldots, \#k_{J+1}\}.$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q(z;z) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q(z;\tilde{z}) \qquad \forall z \in E.$$

Proposition 6.3.2. *The strictly positive measure* $\mathbf{x} := (x(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E)$ *with*

$$x(\mathbf{n},\mathbf{k}) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}(\mathbf{k}), \tag{6.3.2}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{6.3.3}$$

$$\widetilde{\theta}(\mathbf{k}) = \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{\ell=1}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)\ell}} \cdot \phi(\ell)} \cdot \sum_{\tilde{h}=h_{(J+1)\ell}}^{H_{t_{(J+1)\ell}}} b_{t_{(J+1)\ell}}(\tilde{h}) \right), \quad \mathbf{k} \in K, \tag{6.3.4}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z.

Proof. Let us recall some notation:

- It will sometimes be convenient to use the elaborate notation:

$$\mathbf{k} = \Big(\overbrace{k_1, \ldots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}},\ \overbrace{k_{J+1}}^{\substack{\text{central}\\ \text{supplier}}} \Big)$$

$$= \Bigg(\overbrace{\underbrace{[1;\ldots;1]}_{\#k_1 \text{ items}}, \ldots, \underbrace{[J;\ldots;J]}_{\#k_J \text{ items}}}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{\underbrace{\left[t_{(J+1)1}, h_{(J+1)1}; \ldots; h_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}\right]}_{\#k_{J+1} \text{ orders}}}^{\text{central supplier}} \Bigg).$$

- The states of the inventories at the locations $j \in \overline{J}$ are of the form $[j;\ldots;j]$ since there is only raw material of type j at location j.

Note the redundancy of some indicator functions in the global balance equations. We prefer to carry all indicator functions with because it makes it much easier to follow the proof of the stationary distribution.

The global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are given for $(\mathbf{n},\mathbf{k}) \in E$ from (6.3.1) by

flux out of the state $(\mathbf{n},\mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if the inventory at this location is not empty (i.e. $\#k_i > 0$) because of the lost sales rule,
- a service completion of a customer at location $i \in \overline{J}$
  if there is at least one customer (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$),
- a completion of a phase of an order on position ℓ at the central supplier
  if there is at least one order of type $t_{(J+1)\ell}$ (i.e. $\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}$):

$$x\Big(\mathbf{n},\mathbf{k}\Big) \cdot \Bigg(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{\#k_i>0\}} + \sum_{i\in\overline{J}} \mu_i(n_i) \cdot \overbrace{\sum_{\ell=1}^{\#k_{J+1}+1} c(\ell, \#k_{J+1}+1)}^{=1} \cdot \overbrace{\sum_{\tilde{h}=1}^{H_i} b_i(\tilde{h})}^{=1} \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}$$

$$+ \sum_{\ell=1}^{\#k_{J+1}} \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1}) \cdot c(\ell, \#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \Bigg)$$

= flux into the state $(\mathbf{n}, \mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one customer at location i (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$):

$$\sum_{i \in \overline{J}} x\Big(\mathbf{n} - \mathbf{e}_i, \mathbf{k}\Big) \cdot \lambda_i \cdot 1_{\{\#k_i > 0\}} \cdot 1_{\{n_i > 0\}}$$

- a service completion of a customer at location $t_{(J+1)\ell} \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one order at the central supplier (i.e. $\#k_{J+1} > 0$)
  and the order on position ℓ is of type $t_{(J+1)\ell}$ and in phase $h_{(J+1)\ell}$
  (i.e. a customer departs from location $t_{(J+1)\ell}$
  and an item is removed from the associated local inventory there
  and an order is sent to position ℓ in the queue of the central supplier
  and obtain $h_{(J+1)\ell}$ phases):

$$+ \sum_{\ell=1}^{\#k_{J+1}} x\Big(\mathbf{n} + \mathbf{e}_{t_{(J+1)\ell}}, [1; \ldots; 1], \ldots, \underbrace{[t_{(J+1)\ell}; \ldots; t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}+1}, \ldots, [J; \ldots; J],$$
$$\underbrace{\begin{aligned}&[t_{(J+1)1}, h_{(J+1)1}; \ldots; t_{(J+1)(\ell-1)}, h_{(J+1)(\ell-1)}; t_{(J+1)(\ell+1)}, h_{(J+1)(\ell+1)}; \ldots \\ &\qquad \ldots; t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]\end{aligned}}_{\#k_{J+1}-1}\Big)$$
$$\cdot \mu_{t_{(J+1)\ell}}(n_{t_{(J+1)\ell}} + 1) \cdot c(\ell, \#k_{J+1}) \cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}}$$

- phase completion of an order at the central supplier
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one order at the central supplier (i.e. $\#k_{J+1} > 0$)
  and the order on position ℓ is of type $t_{(J+1)\ell}$ and in phase $h_{(J+1)\ell} < H_{t_{(J+1)\ell}}$
  (i.e. the phase of the order on position ℓ is shifted one step down):

$$+ \sum_{\ell=1}^{\#k_{J+1}} x\Big(\mathbf{n}, [1; \ldots; 1], \ldots, [J; \ldots; J],$$
$$\underbrace{\begin{aligned}&[t_{(J+1)1}, h_{(J+1)1}; \ldots; t_{(J+1)(\ell-1)} h_{(J+1)(\ell-1)}; t_{(J+1)\ell}, h_{(J+1)\ell} + 1; t_{(J+1)(\ell+1)}, h_{(J+1)(\ell+1)}; \ldots \\ &\qquad \ldots; t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}]\end{aligned}}_{\#k_{J+1}>0}\Big)$$
$$\cdot \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1}) \cdot c(\ell, \#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} < H_{t_{(J+1)\ell}}\right\}}$$

- a replenishment of the inventory at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one item of raw material at location i (i.e. $\#k_i > 0$) (i.e. an order of type i is in its last phase, i.e. it is removed from the central supplier and is sent to the inventory at location i):

$$+\sum_{i\in\overline{J}} \sum_{\ell=1}^{\#k_{J+1}+1} x\Big(\mathbf{n}, [1;\dots;1], \dots, \underbrace{[i;\dots;i]}_{\#k_i-1}, \dots, [J;\dots;J],$$
$$\underbrace{\big[t_{(J+1)1}, h_{(J+1)1};\dots;t_{(J+1)\ell-1}, h_{(J+1)\ell-1}; i, 1; t_{(J+1)\ell}, h_{(J+1)\ell};\dots;t_{(J+1)\#k_{J+1}}, h_{(J+1)\#k_{J+1}}\big]}_{\#k_{J+1}+1}\Big)$$
$$\cdot \beta_i \cdot \phi(\#k_{J+1}+1) \cdot c(\ell, \#k_{J+1}+1) \cdot 1_{\{\#k_i>0\}}.$$

It has to be shown that the stationary measure (6.3.2) with (6.3.3) and (6.3.4) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

Substitution of (6.3.2) and (6.3.3) into the global balance equations directly leads to

$$\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}$$
$$+\sum_{\ell=1}^{\#k_{J+1}}\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\Big)$$

$$=\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}\setminus\{i\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_i(n_i-1)\cdot\widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot 1_{\{\#k_i>0\}}\cdot 1_{\{n_i>0\}}$$

$$+\sum_{\ell=1}^{\#k_{J+1}}\left(\prod_{j\in\overline{J}\setminus\{t_{(J+1)\ell}\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_{t_{(J+1)\ell}}(n_{t_{(J+1)\ell}}+1)$$
$$\cdot\widetilde{\theta}\Big([1;\ldots;1],\ldots,\underbrace{[t_{(J+1)\ell};\ldots;t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}+1},\ldots,[J;\ldots;J],$$
$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)(\ell-1)},h_{(J+1)(\ell-1)};t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}-1}\Big)$$
$$\cdot\mu_{t_{(J+1)\ell}}(n_{t_{(J+1)\ell}}+1)\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}$$

$$+\sum_{\ell=1}^{\#k_{J+1}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$$
$$\cdot\widetilde{\theta}\Big([1;\ldots;1],\ldots,[J;\ldots;J],$$
$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)(\ell-1)}h_{(J+1)(\ell-1)};t_{(J+1)\ell},h_{(J+1)\ell}+1;t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}>0}\Big)$$
$$\cdot\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\cdot 1_{\left\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\right\}}$$

$$+\sum_{i\in\overline{J}}\sum_{\ell=1}^{\#k_{J+1}+1}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$$
$$\cdot\widetilde{\theta}\Big([1;\ldots;1],\ldots,\underbrace{[i;\ldots;i]}_{\#k_i-1},\ldots,[J;\ldots;J],$$
$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)\ell-1},h_{(J+1)\ell-1};i,1;t_{(J+1)\ell},h_{(J+1)\ell};\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}+1}\Big)$$
$$\cdot\beta_i\cdot\phi(\#k_{J+1}+1)\cdot c(\ell,\#k_{J+1}+1)\cdot 1_{\{\#k_i>0\}}.$$

By substitution of (6.3.3) we obtain

$$\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}$$
$$+\sum_{\ell=1}^{\#k_{J+1}}\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\Big)$$

$$=\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\mu_i(n_i)\cdot 1_{\{\#k_i>0\}}\cdot 1_{\{n_i>0\}}$$

$$+\sum_{\ell=1}^{\#k_{J+1}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$$
$$\cdot\widetilde{\theta}\Big([1;\dots;1],\dots,\underbrace{[t_{(J+1)\ell};\dots;t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}+1},\dots,[J;\dots;J],$$
$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\dots;t_{(J+1)(\ell-1)},h_{(J+1)(\ell-1)};t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\dots\dots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}-1}\Big)$$
$$\cdot\lambda_{t_{(J+1)\ell}}\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}$$

$$+\sum_{\ell=1}^{\#k_{J+1}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$$
$$\cdot\widetilde{\theta}\Big([1;\dots;1],\dots,[J;\dots;J],$$
$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\dots;t_{(J+1)(\ell-1)}h_{(J+1)(\ell-1)};t_{(J+1)\ell},h_{(J+1)\ell}+1;t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\dots\dots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}>0}\Big)$$
$$\cdot\beta_{(J+1)\ell}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\cdot 1_{\left\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\right\}}$$

$$+\sum_{i\in\overline{J}}\sum_{\ell=1}^{\#k_{J+1}+1}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$$
$$\cdot\widetilde{\theta}\Big([1;\dots;1],\dots,\underbrace{[i;\dots;i]}_{\#k_i-1},\dots,[J;\dots;J],$$
$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\dots;t_{(J+1)\ell-1},h_{(J+1)\ell-1};i,1;t_{(J+1)\ell},h_{(J+1)\ell};\dots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}+1}\Big)$$
$$\cdot\beta_i\cdot\phi(\#k_{J+1}+1)\cdot c(\ell,\#k_{J+1}+1)\cdot 1_{\{\#k_i>0\}}.$$

Cancelling $\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$ and the sums with the terms $\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}$ on both sides of the equation leads to

$$\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{\ell=1}^{\#k_{J+1}}\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\Big)$$

$$\begin{aligned}
&=\sum_{\ell=1}^{\#k_{J+1}}\widetilde{\theta}\Big([1;\ldots;1],\ldots,\underbrace{[t_{(J+1)\ell};\ldots;t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}+1},\ldots,[J;\ldots;J],\\
&\quad\underbrace{\big[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)(\ell-1)},h_{(J+1)(\ell-1)};t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\ldots\;\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}\big]}_{\#k_{J+1}-1}\Big)\\
&\quad\cdot\lambda_{t_{(J+1)\ell}}\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\\
&\quad+\sum_{\ell=1}^{\#k_{J+1}}\widetilde{\theta}\Big([1;\ldots;1],\ldots,[J;\ldots;J],\\
&\quad\underbrace{\big[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)(\ell-1)}h_{(J+1)(\ell-1)};t_{(J+1)\ell},h_{(J+1)\ell}+1;t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\ldots\;\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}\big]}_{\#k_{J+1}>0}\Big)\\
&\quad\cdot\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\cdot 1_{\left\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\right\}}\\
&\quad+\sum_{i\in\overline{J}}\sum_{\ell=1}^{\#k_{J+1}+1}\widetilde{\theta}\Big([1;\ldots;1],\ldots,\underbrace{[i;\ldots;i]}_{\#k_i-1},\ldots,[J;\ldots;J],\\
&\quad\underbrace{\big[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)\ell-1},h_{(J+1)\ell-1};i,1;t_{(J+1)\ell},h_{(J+1)\ell};\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}\big]}_{\#k_{J+1}+1}\Big)\\
&\quad\cdot\beta_i\cdot\phi(\#k_{J+1}+1)\cdot c(\ell,\#k_{J+1}+1)\cdot 1_{\{\#k_i>0\}}.
\end{aligned}\tag{6.3.5}$$

By substitution of (6.3.4) we obtain for the third summand on the right side of (6.3.5)

$$\sum_{i\in\overline{J}}\sum_{\ell=1}^{\#k_{J+1}+1}\widetilde{\theta}\Big([1;\ldots;1],\ldots,\underbrace{[i;\ldots;i]}_{\#k_i-1},\ldots,[J;\ldots;J],$$
$$\underbrace{\left[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)\ell-1},h_{(J+1)\ell-1};i,1;t_{(J+1)\ell},h_{(J+1)\ell};\ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}\right]}_{\#k_{J+1}+1}\Big)$$
$$\cdot\beta_i\cdot\phi(\#k_{J+1}+1)\cdot c(\ell,\#k_{J+1}+1)\cdot 1_{\{\#k_i>0\}}$$

$$=\sum_{i\in\overline{J}}\sum_{\ell=1}^{\#k_{J+1}+1}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\lambda_i\cdot\prod_{m=1}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot\left(\frac{1}{\beta_i\cdot\phi(\#k_{J+1}+1)}\cdot\sum_{\tilde{h}=1}^{H_i}b_i(\tilde{h})\right)\cdot\beta_i\cdot\phi(\#k_{J+1}+1)\cdot c(\ell,\#k_{J+1}+1)\cdot 1_{\{\#k_i>0\}}$$

$$=\sum_{i\in\overline{J}}\underbrace{\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{m=1}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)}_{=\widetilde{\theta}(\mathbf{k})}\cdot\underbrace{\left(\sum_{\tilde{h}=1}^{H_i}b_i(\tilde{h})\right)}_{=1}$$
$$\cdot\underbrace{\sum_{\ell=1}^{\#k_{J+1}+1}c(\ell,\#k_{J+1}+1)}_{=1}\cdot\lambda_i\cdot 1_{\{\#k_i>0\}}$$

$$=\sum_{i\in\overline{J}}\widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot 1_{\{\#k_i>0\}}. \qquad (6.3.6)$$

By substitution of (6.3.4) we obtain for the second summand on the right side of (6.3.5)

$$\sum_{\ell=1}^{\#k_{J+1}} \widetilde{\theta}\Big([1;\dots;1],\dots,[J;\dots;J],$$
$$\underbrace{\big[t_{(J+1)1},h_{(J+1)1};\dots;t_{(J+1)(\ell-1)}h_{(J+1)(\ell-1)};t_{(J+1)\ell},h_{(J+1)\ell}+1;t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\dots \\ \dots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}\big]}_{\#k_{J+1}>0}\Big)$$
$$\cdot\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\cdot 1_{\left\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\right\}}$$

$$=\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot\left(\frac{1}{\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})}\cdot\sum_{\tilde{h}=h_{(J+1)\ell}+1}^{H_{t_{(J+1)\ell}}}b_{t_{(J+1)\ell}}(\tilde{h})\right)$$
$$\cdot\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\cdot 1_{\left\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\right\}}$$

$$=\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)\cdot\left(\sum_{\tilde{h}=h_{(J+1)\ell}+1}^{H_{t_{(J+1)\ell}}}b_{t_{(J+1)\ell}}(\tilde{h})\right)$$
$$\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}}\cdot 1_{\left\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\right\}}. \tag{6.3.7}$$

By substitution of (6.3.4) we obtain for the first summand on the right side of (6.3.5)

$$\sum_{\ell=1}^{\#k_{J+1}} \widetilde{\theta}\Big([1;\ldots;1],\ldots,\underbrace{[t_{(J+1)\ell};\ldots;t_{(J+1)\ell}]}_{\#k_{t_{(J+1)\ell}}+1},\ldots,[J;\ldots;J],$$

$$\underbrace{[t_{(J+1)1},h_{(J+1)1};\ldots;t_{(J+1)(\ell-1)},h_{(J+1)(\ell-1)};t_{(J+1)(\ell+1)},h_{(J+1)(\ell+1)};\ldots \ldots;t_{(J+1)\#k_{J+1}},h_{(J+1)\#k_{J+1}}]}_{\#k_{J+1}-1}\Big)$$

$$\cdot\lambda_{t_{(J+1)\ell}}\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}$$

$$=\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\frac{1}{\lambda_{t_{(J+1)\ell}}}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{(J+1)t_m}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot\lambda_{t_{(J+1)\ell}}\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}$$

$$=\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}$$

$$=\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}\cdot 1_{\{h_{(J+1)\ell}=H_{t_{(J+1)\ell}}\}}$$
$$+\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}\cdot 1_{\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\}}$$

$$=\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{m=1}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)\cdot\beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})$$
$$\cdot c(\ell,\#k_{J+1})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}\cdot 1_{\{h_{(J+1)\ell}=H_{t_{(J+1)\ell}}\}}$$
$$+\sum_{\ell=1}^{\#k_{J+1}}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\substack{m=1\\m\neq\ell}}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)m}}\cdot\phi(m)}\cdot\sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}}b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot c(\ell,\#k_{J+1})\cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell})\cdot 1_{\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\}}\cdot 1_{\{h_{(J+1)\ell}<H_{t_{(J+1)\ell}}\}}. \tag{6.3.8}$$

Addition of (6.3.8) and (6.3.7) yields

$$\sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m=1}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right) \cdot \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1})$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} = H_{t_{(J+1)\ell}}\right\}}$$
$$+ \sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{\substack{m=1\\ m\neq\ell}}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right)$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot b_{t_{(J+1)\ell}}(h_{(J+1)\ell}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} < H_{t_{(J+1)\ell}}\right\}}$$
$$+ \sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{\substack{m=1\\ m\neq\ell}}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right) \cdot \sum_{\tilde{h}=h_{(J+1)\ell}+1}^{H_{t_{(J+1)\ell}}} b_{t_{(J+1)\ell}}(\tilde{h})$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} < H_{t_{(J+1)\ell}}\right\}}$$

$$= \sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m=1}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right) \cdot \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1})$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} = H_{t_{(J+1)\ell}}\right\}}$$
$$+ \sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{\substack{m=1\\ m\neq\ell}}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right) \cdot \sum_{\tilde{h}=h_{(J+1)\ell}}^{H_{t_{(J+1)\ell}}} b_{t_{(J+1)\ell}}(\tilde{h})$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} < H_{t_{(J+1)\ell}}\right\}}$$

$$= \sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m=1}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right) \cdot \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1})$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} = H_{t_{(J+1)\ell}}\right\}}$$
$$+ \sum_{\ell=1}^{\#k_{J+1}} \prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m=1}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right) \cdot \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1})$$
$$\cdot\, c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}} \cdot 1_{\left\{h_{(J+1)\ell} < H_{t_{(J+1)\ell}}\right\}}$$

$$= \sum_{\ell=1}^{\#k_{J+1}} \underbrace{\prod_{j\in\bar{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{m=1}^{\#k_{J+1}} \left(\frac{1}{\beta_{t_{(J+1)m}} \cdot \phi(m)} \cdot \sum_{\tilde{h}=h_{(J+1)m}}^{H_{t_{(J+1)m}}} b_{t_{(J+1)m}}(\tilde{h})\right)}_{=\tilde{\theta}(\mathbf{k})}$$
$$\cdot\, \beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1}) \cdot c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}}$$

$$= \sum_{\ell=1}^{\#k_{J+1}} \tilde{\theta}(\mathbf{k}) \cdot \left(\beta_{t_{(J+1)\ell}} \cdot \phi(\#k_{J+1}) \cdot c(\ell,\#k_{J+1}) \cdot 1_{\left\{\#k_{t_{(J+1)\ell}} < b_{t_{(J+1)\ell}}\right\}}\right). \tag{6.3.9}$$

Consequently, because of (6.3.6) and (6.3.9) the right-hand side of (6.3.5) is

$$\sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot 1_{\{\#k_i>0\}} + \sum_{\ell=1}^{\#k_{J+1}} \widetilde{\theta}(\mathbf{k}) \sum_{\ell=1}^{\#k_{J+1}} \beta_{t_{(J+1)\ell}}\cdot\phi(\#k_{J+1})\cdot c(\ell,\#k_{J+1})\cdot 1_{\left\{\#k_{t_{(J+1)\ell}}<b_{t_{(J+1)\ell}}\right\}},$$

which is obviously the left-hand side.

Inspection of the system (6.3.5) reveals that it is a "generator equation", i.e. the global balance equation $\widetilde{\theta}\cdot\mathbf{Q}_{red}=\mathbf{0}$ for a suitably defined ergodic Markov process on state space K with "reduced generator" $\mathbf{Q}_{red}$. Because the Markov process generated by $\mathbf{Q}_{red}$ is irreducible the solution of (6.3.5) is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 6.3.3. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k})\in E), \quad \pi(\mathbf{n},\mathbf{k}) := \lim_{t\to\infty} P(Z(t)=(\mathbf{n},\mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := \left(\xi(\mathbf{n}) : \mathbf{n}\in\mathbb{N}_0^J\right), \quad \xi(\mathbf{n}) := \lim_{t\to\infty} P((X_1(t),\dots,X_J(t))=\mathbf{n}),$$

$$\theta := (\theta(\mathbf{k}) : \mathbf{k}\in K), \quad \theta(\mathbf{k}) := \lim_{t\to\infty} P((Y_1(t),\dots,Y_J(t),W_{J+1}(t))=\mathbf{k}).$$

Theorem 6.3.4. *The queueing-inventory process Z is ergodic if and only if for $j\in\overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}<\infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n},\mathbf{k}) = \xi(\mathbf{n})\cdot\theta(\mathbf{k}), \tag{6.3.10}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}}\xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}, \quad n_j\in\mathbb{N}_0,\ j\in\overline{J}, \tag{6.3.11}$$

$$\theta(\mathbf{k}) = C_\theta^{-1}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\ell=1}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)\ell}}\cdot\phi(\ell)}\cdot\sum_{\tilde{h}=h_{(J+1)\ell}}^{H_{t_{(J+1)\ell}}} b_{t_{(J+1)\ell}}(\tilde{h})\right), \quad \mathbf{k}\in K, \tag{6.3.12}$$

and normalisation constants

$$C_j = \sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)},$$

$$C_\theta = \sum_{\mathbf{k}\in K}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\ell=1}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)\ell}}\cdot\phi(\ell)}\cdot\sum_{\tilde{h}=h_\ell}^{H_{t_{(J+1)\ell}}} b_{t_{(J+1)\ell}}(\tilde{h})\right).$$

Proof. Z is ergodic, if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x}\cdot\mathbf{Q}=\mathbf{0}$ from Proposition 6.3.2 can be normalised (i.e. $\sum_{\mathbf{n}\in\mathbb{N}_0}\sum_{\mathbf{k}\in K}x(\mathbf{n},\mathbf{k})<\infty$). Because of Proposition 6.3.2 it holds

$$\begin{aligned}\sum_{\mathbf{n}\in\mathbb{N}_0}\sum_{\mathbf{k}\in K}x(\mathbf{n},\mathbf{k}) &= \sum_{\mathbf{n}\in\mathbb{N}_0}\widetilde{\xi}(\mathbf{n})\cdot\sum_{\mathbf{k}\in K}\widetilde{\theta}(\mathbf{k})\\ &= \left(\prod_{j\in\overline{J}}\sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}\right)\\ &\quad\cdot\left(\sum_{\mathbf{k}\in K}\prod_{j\in\overline{J}}\left(\frac{1}{\lambda_j}\right)^{\#k_j}\cdot\prod_{\ell=1}^{\#k_{J+1}}\left(\frac{1}{\beta_{t_{(J+1)\ell}}\cdot\phi(\ell)}\cdot\sum_{\tilde{h}=h_{(J+1)\ell}}^{H_{t_{(J+1)\ell}}}b_{t_{(J+1)\ell}}(\tilde{h})\right)\right).\end{aligned}$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 6.3.2 can be normalised if and only if $\sum_{n_j\in\mathbb{N}_0}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}<\infty$ for all $j\in\overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n},\mathbf{k})=\frac{x(\mathbf{n},\mathbf{k})}{\sum_{\mathbf{n}\in\mathbb{N}_0}\sum_{\mathbf{k}\in K}x(\mathbf{n},\mathbf{k})},$$

where $x(\mathbf{n},\mathbf{k})$ is given in Proposition 6.3.2. □

Remark 6.3.5. The expression (6.3.10) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

The explicit formula (6.3.12) for θ shows that in fact there exists a three-term product structure, and that moreover the equilibrium for the integrated model is stratified. In the upper stratum, we have two independent vectors for production and inventory-replenishment, the latter splits into two products, a factor for the subsystem comprising the inventories and a factor for the replenishment subsystem.

In the lower stratum, each of the three factors of the upper stratum is decomposed completely in "single-component" factors concerning the production servers, the inventories, and the replenishment server. It should be noted that the factors for the inventories and the replenishment server do not indicate internal independence, but they are of product form as the celebrated conditionally independent coordinates in the equilibrium of Gordon-Newell networks (see Theorem A.2.6).

Representation (6.3.11) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(\mathbf{n})=\prod_{j\in\overline{J}}\xi_j(n_j)=\prod_{j\in\overline{J}}C_j^{-1}\prod_{\ell=1}^{n_j}\frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. In our production network, however, servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (6.3.10) has been unexpected to us.

Our production-inventory-replenishment system can be considered as a "Jackson network in a random environment" in [KDO16, Section 4]. We can interpret the inventory-replenishment subsystem, which contributes via θ to Theorem 6.3.4, as a "random environment" for the production network of nodes $\overline{J}$, which is a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1], we conclude from the hindsight that decoupling of the queueing process $(X_1, \ldots, X_J)$ and the process $(Y_1, \ldots, Y_J, W_{J+1})$, i.e. the formula (6.3.10), is a consequence of that Theorem 4.1. It should be noted that the three-term structure of the upper stratum of the product form steady state π in Theorem 6.3.4 can not be obtained from the general theory.

Our direct proof of Theorem 6.3.4 is much shorter than embedding the present model into the general framework of [KDO16].
The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

Note that in (6.3.12) the term

$$\frac{1}{\beta_{t_{(J+1)\ell}}} \cdot \sum_{\tilde{h}=h_{(J+1)\ell}}^{H_{t_{(J+1)\ell}}} b_{t_{(J+1)\ell}}(\tilde{h})$$

represents the stationary residual life time distribution in a stationary renewal process with life time distribution $B_{t_{(J+1)\ell}}$.

6.4. Exponentially distributed service time

As mentioned before by varying the parameters of the phase-type distribution any distribution on $(\mathbb{R}_0^+, \mathbb{B}_0^+)$ can be approximated sufficiently close. A phase-type distribution with one phase is an exponential distribution. In this section, we present the results for the queueing-inventory system where the service time of the central supplier is exponentially distributed with type-dependent rate ν_j, $j \in \overline{J}$, as described in Section 6.2. Furthermore, the structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the contents of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_{J+1}(t)$ we denote the sequence of orders at the central supplier at time $t \geq 0$.

We denote by K_{J+1} the set of possible states at the central supplier (local state space). The state

$$k_{J+1} := \left[t_{(J+1)1}; \ldots ; t_{(J+1)\#k_{J+1}}\right] \in K_{J+1}$$

indicates that there are $\#k_{J+1}$ orders at the supplier and on position $p \in \{1, \ldots, \#k_{J+1}\}$ resides an order of type $t_{(J+1)p} \in \overline{J}$. More precisely, $t_{(J+1)1}$ is the order at the head of the line, that is in service and $t_{(J+1)\#k_{J+1}}$ is the order at the tail of the line.

Notational convention. In order to make the reading easier, we will use a unified notation for the states of the inventories at the locations and the state of the workstation at the central supplier. Hence, if at time $t \geq 0$ the inventory size of $j \in \overline{J}$ is $Y_j(t) = \#k_j$, then the blown up state of the inventory at location $j \in \overline{J}$ is

$$k_j = \underbrace{[j; \ldots ; j]}_{\#k_j \text{ items}} \in K_j .$$

The global states of the inventory-replenishment subsystem are then

$$\mathbf{k} = \Big(\overbrace{k_1, \ldots, k_J,}^{\substack{\text{inventories} \\ \text{at locations}}} \overbrace{k_{J+1}}^{\substack{\text{central} \\ \text{supplier}}}\Big) \in K \subseteq \prod_{j=1}^{J+1} K_j ,$$

where K_j denotes the local state space at j and K denotes the set of feasible states composed of feasible local states.

For $\#k_j = 0$, $j \in \overline{J}$, we read

$$[j; \ldots ; j] =: [0]$$

and for $\#k_{J+1} = 0$ we read

$$\left[t_{(J+1)1}; \ldots ; t_{(J+1)\#k_{(J+1)}}\right] =: [0].$$

We define the joint queueing-inventory process of this system by

$$Z = \left((X_1(t), \ldots, X_J(t), Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) \,:\, t \geq 0\right).$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \left\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J, \mathbf{k} \in K\right\}.$$

Definition 6.4.1. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi(\mathbf{n}, \mathbf{k}) : (\mathbf{n}, \mathbf{k}) \in E), \quad \pi(\mathbf{n}, \mathbf{k}) := \lim_{t\to\infty} P(Z(t) = (\mathbf{n}, \mathbf{k}))$$

and the appropriate marginal distributions

$$\xi := \left(\xi(\mathbf{n}) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi(\mathbf{n}) := \lim_{t\to\infty} P((X_1(t), \ldots, X_J(t)) = \mathbf{n}),$$

$$\theta := (\theta(\mathbf{k}) : \mathbf{k} \in K), \quad \theta(\mathbf{k}) := \lim_{t\to\infty} P((Y_1(t), \ldots, Y_J(t), W_{J+1}(t)) = \mathbf{k}).$$

Recall that the system is irreducible and regular. This leads to our main result, which is a special case of Theorem 6.3.4.

Theorem 6.4.2. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n}, \mathbf{k}) = \xi(\mathbf{n}) \cdot \theta(\mathbf{k}), \tag{6.4.1}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{6.4.2}$$

$$\theta(\mathbf{k}) = C_\theta^{-1} \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{\ell=1}^{\#k_{J+1}} \frac{1}{\nu_{t_{(J+1)\ell}} \cdot \phi(\ell)}, \quad \mathbf{k} \in K, \tag{6.4.3}$$

and normalisation constants

$$C_j = \sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} \quad \text{and} \quad C_\theta = \sum_{\mathbf{k}\in K} \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{\#k_j} \cdot \prod_{\ell=1}^{\#k_{J+1}} \frac{1}{\nu_{t_{(J+1)\ell}} \cdot \phi(\ell)}.$$

7. Production-inventory system with (r_j, S_j)-policy

In Chapter 2 to Chapter 6 we have focused on base stock policies. In classic inventory theory, several replenishment policies are considered. In this chapter, we look at the (r, S)-policy. The (r, S)-policy means: If the size of the local inventory is less than or equal to the reorder level $r \geq 0$, a replenishment order is placed instantaneously. With each replenishment the local inventory level is restocked to exactly $S < \infty$. The maximal size of the inventory is S. We assume that $0 < S$ and that there is at most one outstanding order ($r < S$). A brief discussion of the advantages and disadvantages can be found in [SPP98, pp. 238f.] for (r, S)-policy.

7.1. Related literature and own contributions

There are some papers on queueing-inventory systems with (r, S)-policy as we see in the literature review on page viii, so we repeat only the most relevant sources for our investigation. For a survey paper we refer to Krishnamoorthy et al. [KLM11].

The most relevant source for our present investigations is the article of Schwarz and her coauthors [SSD$^+$06]. They investigate $M/M/1$ systems with inventory management, exponentially distributed lead times and lost sales. They consider among other replenishment policies the (r, S)-policy. Further, they distinguish between an infinite and a finite waiting room. They derive stationary distributions of joint queue length and inventory processes in explicit product form and calculate performance measures of the respective systems.

Krishnamoorthy and Viswanath study in [KV13] production-inventory systems with (r, S)-policy, positive service time and lost sales. They derive the joint stationary distribution in explicit product form. They develop a technique where the steady state vector of the classical $M/M/1$ system and the steady state vector of a production-inventory system, where the service is instantaneous and no backlogs are allowed, are combined. They apply their technique to the models discussed in [SSD$^+$06].

Our main contributions are the following:
The systems under investigation differ with respect to the reorder level and the number of locations and workstations. In Section 7.3 we analyse a system with $(0, S_j)$-policy with J locations a supplier network consisting of M workstations and in Section 7.4 a system with $(1, S_j)$-policy with two locations and a supplier network consisting of one workstation. For these systems we develop Markov process models and derive their stationary distributions of the joint queueing-inventory processes in explicit product form. A cost analysis can be performed as for the basic model in Section 5.4 on page 137.

The articles in the literature about production-inventory systems with (r, S)-policy are only single location models. Our results extend their settings to a multi-dimensional system. Furthermore, the lead time of an order is composed of the waiting times plus the service times in the supplier network. They are therefore more complex than exponential lead times considered in the literature.

7.2. Description of the general model

The supply chain of interest is depicted in Figure 7.2.1. We have a set of locations $\overline{J} := \{1, 2, \ldots, J\}$. Each of the locations consists of a production system with an attached inventory. The inventories are replenished by a supplier network, which consists of a set of workstations $\overline{M} := \{J + 1, \ldots, J + M\}$ and manufactures raw material for all locations, but distinguishes between the replenishment orders from different locations. Each order of raw material is specified by a location $j \in \overline{J}$ and the resulting raw material is sent back to the location that has placed the order.

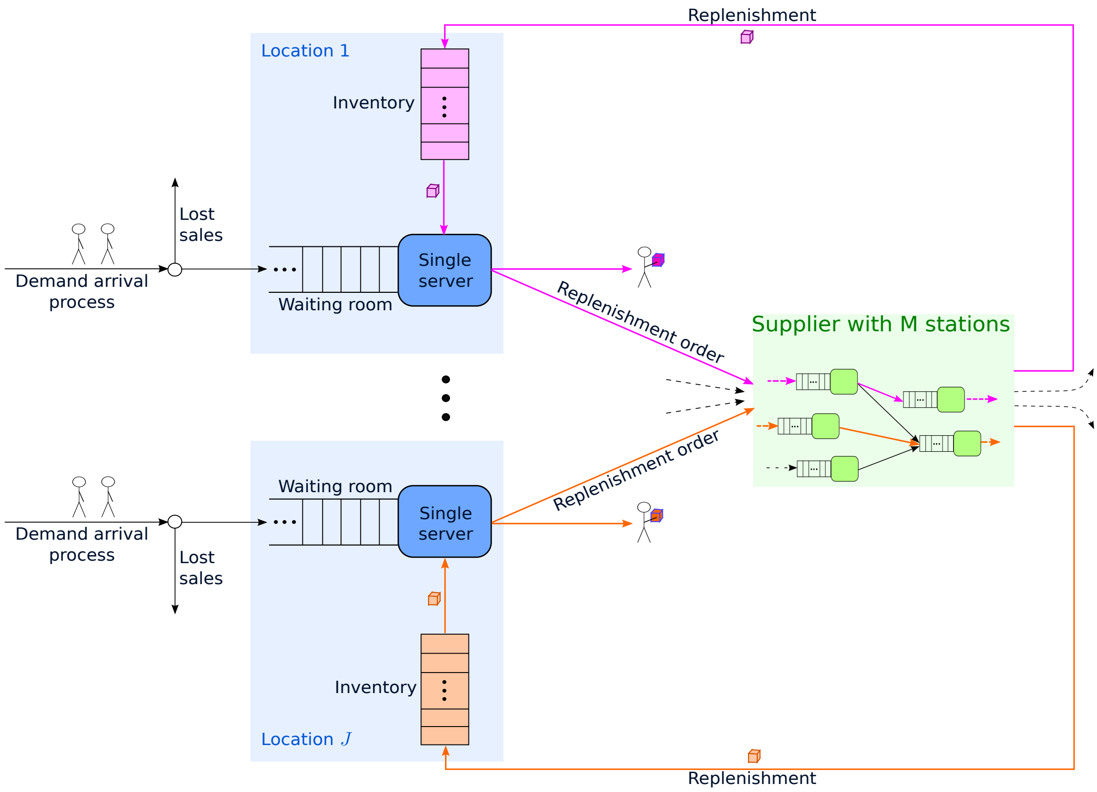


Figure 7.2.1.: Supply chain with (r_j, S_j)-policy at location j

Facilities in the supply chain. Each production system $j \in \overline{J}$ consists of a single server (machine) with infinite waiting room that serves customers on a make-to-order basis under a FCFS regime. Customers arrive one by one at production system j according to a Poisson process with rate $\lambda_j > 0$ and require service. To satisfy a customer's demand the production system needs exactly one item of raw material, which is taken from the associated local inventory. When a new customer arrives at a location while the previous customer's order is not finished, this customer will wait. If the inventory is depleted at location j, the customers who are already waiting in line will wait, but new arriving customers at this location will decide not to join the queue and are lost ("local lost sales").

The service requests at the locations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at location $j \in \overline{J}$ is provided with local queue-length-dependent intensity. If there are $n_j > 0$ customers present at location j either waiting or in service (if any) and if the inventory is not depleted, the service intensity is $\mu_j(n_j) > 0$. If the server is ready to serve a customer who is at the head of the line and the inventory is not depleted, the service immediately starts. Otherwise, the service starts at the instant of time when the next replenishment arrives at the local inventory.

The inventory at location $j \in \overline{J}$ is controlled by the (r_j, S_j)-policy. This means, if the size of the local inventory is less than or equal to the reorder level r_j, a replenishment order is instantaneously placed. With each replenishment the local inventory level is restocked to exactly $S_j < \infty$. The maximal size of the inventory at location j is S_j. We assume that $S_j > 0$ and that there is at most one outstanding order.

The systems under investigation differ with respect to the reorder level and the number of locations and the number of workstations in the supplier network in the following way:

- $(0, S_j)$-policy with J locations and M workstations (Section 7.3),
- $(1, S_j)$-policy with 2 locations and one workstation (Section 7.4).

Routing in the supply chain. A customer departs from the system (with the consumed material) immediately after the service.

To distinguish orders from different locations, we mark each order by a "type", which for simplicity is the index of the location where the order is triggered.

An order triggered by location $j \in \overline{J}$ follows a type-j-dependent route for eventual replenishment, denoted by $r(j) = (r(j,1), \ldots, r(j, S(j)-1), r(j, S(j)))$, where $r(j,\ell) \in \overline{M}$ for $\ell = 2, \ldots, S(j)$ is the identifier of the ℓ-th workstation on the path $r(j)$, and $S(j)$ is the number of stages of the route of type j. For completeness we fix $r(j,1) := j \in \overline{J}$.

The workstation $m \in \overline{M}$ consists of a single server and a waiting room under a FCFS regime. The service requests at the workstations are exponentially-1 distributed. All service requests constitute an independent family of random variables which are independent of the arrival streams. The service at workstation m is provided with local queue-length-dependent intensity. If there are $\ell > 0$ orders present, the service intensity is $\nu_m(\ell) > 0$.

It is assumed that transmission times for orders are negligible and set to zero and that transportation times between the supplier network and the inventory are negligible.

The usual independence assumptions are assumed to hold as well.

Remark 7.2.1. With respect to economic aspects, if at all workstations $m \in \overline{M}$ it holds $\nu_m(1) = \infty$ ($\equiv$ "lead time is zero" assumption), a reorder level $r_j > 0$ at all locations $j \in \overline{J}$ under any policy does not make sense, since r_j items of raw material are never touched by the customers and remain in stock forever (cf. [SSD+06, p. 63]).

Remark 7.2.2. These are standard notations. Be careful not to confuse the notations S_j and $S(j)$ as well as r_j and $r(j)$. r_j and S_j are parameters from the (r_j, S_j)-policy of location $j \in \overline{J}$, more precisely r_j is the reorder level and S_j is the order-up-to level. $r(j)$ is the type-j-dependent route (path) for eventual replenishment and $S(j)$ is the number of states of the route of type $j \in \overline{J}$. These notations are standard in the literature on inventory theory on one side and queueing theory on the other side.

7.3. $(0, S_j)$-policy with J locations and M workstations

In this section, we study the production-inventory system as described in Section 7.2, where the reorder level r_j at location $j \in \overline{J}$ is equal to zero. Therefore, if the local inventory at location j is depleted, a replenishment order is instantaneously placed. With each replenishment the local inventory level is restocked to exactly $S_j < \infty$. We assume that $0 < S_j$ and that there is at most one outstanding order. We denote $\mathbf{S} := \big(S_j : j \in \overline{J}\big)$.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the contents of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_m(t)$ we denote the sequence of orders at workstation $m \in \overline{M}$ of the supplier network at time $t \geq 0$.

We denote by K_m the set of possible states at node $m \in \overline{M}$ (local state space).
The state

$$k_m := \underbrace{[t_{m1}, s_{m1}; \ldots ; t_{m\#k_m}, s_{m\#k_m}]}_{\#k_m \text{ orders}} \in K_m$$

indicates that there are $\#k_m$ orders at workstation $m \in \overline{M}$, on position $p \in \{1, \ldots, \#k_m\}$ resides an order of type $t_{mp} \in \overline{J}$, which is on stage $s_{mp} \in \{1, \ldots, S(t_{mp})\}$ of its route $r(t_{mp}) = \Big(r(t_{mp}, 1), \ldots, r(t_{mp}, S(t_{mp}))\Big)$. More precisely, (t_{m1}, s_{m1}) is the order at the head of the line, which is in service, and $(t_{m\#k_m}, s_{m\#k_m})$ is the order at the tail of the line.

Notational convention. To improve readability, we use a unified notation for the states of the inventories at the locations and the states of the workstations in the supplier network. In doing this we identify items of raw material arriving at the inventory $j \in \overline{J}$ with the order sent out to the supplier network when an item is consumed by a departing customer. Therefore, adopting the state description of the workstations for that of the inventories, the blown up state of the inventory at location $j \in \overline{J}$ at time $t \geq 0$ is

$$Y_j(t) = k_j = \underbrace{[j, 1; \ldots ; j, 1]}_{\#k_j \text{ items}},$$

since the route of type j starts in the inventory at location j (i.e. $s_{jp} = 1$ for all $p \in \{1, \ldots, \#k_j\}$). A stage number $s_{jp} > 1$ indicates that the unit (as an order) is in the supplier network.

The global states of the inventory-replenishment subsystem are then

$$\mathbf{k} = \Big(\overbrace{k_1, \dots, k_J}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{k_{J+1}, \dots, k_{J+M}}^{\text{supplier network}}\Big) \in K \subseteq \prod_{j=1}^{J+M} K_j,$$

where K_j denotes the local state space at $j \in \overline{J} \cup \overline{M}$ and K denotes the feasible states composed of feasible local states.

For $\#k_j = 0$, $j \in \overline{J}$, we read

$$\big[t_{j1}, s_{j1}; \dots; t_{j\#k_j}, s_{j\#k_j}\big] =: [0],$$

and for $\#k_m = 0$, $m \in \overline{M}$, we read

$$[t_{m1}, s_{m1}; \dots; t_{m\#k_m}, s_{m\#k_m}] =: [0].$$

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), \dots, X_J(t), Y_1(t), \dots, Y_J(t), W_{J+1}(t), \dots, W_{J+M}(t)) \,:\, t \geq 0)\,.$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \big\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J, \, \mathbf{k} \in K\big\}\,.$$

7.3.1. Limiting and stationary behaviour

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z; \tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n}, \mathbf{k}) \in E$, where a typical state is (we will impose necessary restrictions if needed)

$$\begin{aligned}(\mathbf{n}, \mathbf{k}) = \Big(\mathbf{n}, &\underbrace{[1, 1; \dots; 1, 1]}_{\#k_1}, \dots, \underbrace{[J, 1; \dots; J, 1]}_{\#k_J},\\ &\big[t_{(J+1)1}, s_{(J+1)1}; \dots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\big], \dots\\ &\dots, \big[t_{(J+M)1}, s_{(J+M)1}; \dots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\big]\Big).\end{aligned} \tag{7.3.1}$$

- Arrival of a customer at location $i \in \overline{J}$,
  which happens only if the inventory at this location is not empty because of the lost sales rule:
  $$q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n}+\mathbf{e}_i, \mathbf{k})\right) = \lambda_i \cdot 1_{\{\#k_i > 0\}}, \qquad i \in \overline{J}.$$

- Service completion of a customer at location $i \in \overline{J}$, which happens only if there is at least one customer at location i and either

- if there is one item of raw material present at location i (i.e. $\#k_i = 1$), i.e. from location i (= station $r(i,1)$), where $\#k_i = 1$ items are present, a customer departs and an item of raw material is removed from the associated local inventory, in addition a replenishment order is sent to workstation $r(i,2) \in \overline{M}$ of the supplier network, where $\#k_{r(i,2)}$ units (orders) have already been present:

$$q\Bigg(\Big(\mathbf{n}, \underbrace{[1,1;\ldots;1,1]}_{\#k_1}, \ldots, \underbrace{[i,1]}_{\#k_i=1}, \ldots, \underbrace{[J,1;\ldots;J,1]}_{\#k_J}, \\ \big[t_{(J+1)1}, s_{(J+1)1}; \ldots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\big], \ldots \\ \ldots, \underbrace{\Big[t_{r(i,2)1}, s_{r(i,2)1}; \ldots; t_{r(i,2)\#k_{r(i,2)}}, s_{r(i,2)\#k_{r(i,2)}}\Big]}_{\#k_{r(i,2)}}, \ldots \\ \ldots, \big[t_{(J+M)1}, s_{(J+M)1}; \ldots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\big]\Big); \\ \Big(\mathbf{n}-\mathbf{e}_i, \underbrace{[1,1;\ldots;1,1]}_{\#k_1}, \ldots, \underbrace{[0]}_{\#k_i-1=0}, \ldots, \underbrace{[J,1;\ldots;J,1]}_{\#k_J}, \\ \big[t_{(J+1)1}, s_{(J+1)1}; \ldots; t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\big], \ldots \\ \ldots, \underbrace{\Big[t_{r(i,2)1}, s_{r(i,2)1}; \ldots; t_{r(i,2)\#k_{r(i,2)}}, s_{r(i,2)\#k_{r(i,2)}}; i, 2\Big]}_{\#k_{r(i,2)}+1}, \ldots \\ \ldots, \big[t_{(J+M)1}, s_{(J+M)1}; \ldots; t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\big]\Big)\Bigg) \\ = \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i=1\}}, \qquad i \in \overline{J}. \tag{7.3.2}$$

Notational convention: For transition rates like the one above we will henceforth use an abbreviated notation. Using this abbreviation (7.3.2) reads

$$q\Bigg(\Big(\mathbf{n}, \ldots, \underbrace{[i,1]}_{\#k_i=1}, \ldots, \underbrace{\Big[t_{r(i,2)1}, s_{r(i,2)1}; \ldots; t_{r(i,2)\#k_{r(i,2)}}, s_{r(i,2)\#k_{r(i,2)}}\Big]}_{\#k_{r(i,2)}}, \ldots\Big); \\ \Big(\mathbf{n}-\mathbf{e}_i, \ldots, \underbrace{[0]}_{\#k_i-1=0}, \ldots, \underbrace{\Big[t_{r(i,2)1}, s_{r(i,2)1}; \ldots; t_{r(i,2)\#k_{r(i,2)}}, s_{r(i,2)\#k_{r(i,2)}}; i, 2\Big]}_{\#k_{r(i,2)}+1}, \ldots\Big)\Bigg) \\ = \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i=1\}}, \qquad i \in \overline{J}.$$

This means that we will explicitly write only those local states of $(\mathbf{n}, \mathbf{k})$ and its successor state that are relevant for the described transition.

- or if $\#k_i > 1$ items of raw material are present at location i,
  i.e. from location i (= station $r(i,1)$), where $\#k_i > 1$ items are present, a customer departs and an item of raw material is removed from the associated local inventory:

$$q\Bigg(\Big(\mathbf{n},\dots,\underbrace{[i,1;\dots;i,1]}_{\#k_i>1},\dots,\underbrace{\Big[t_{r(i,2)1},s_{r(i,2)1};\dots;t_{r(i,2)\#k_{r(i,2)}},s_{r(i,2)\#k_{r(i,2)}}\Big]}_{\#k_{r(i,2)}},\dots\Big);$$
$$\Big(\mathbf{n}-\mathbf{e}_i,\dots,\underbrace{[i,1;\dots;i,1]}_{\#k_i-1>0},\dots,\underbrace{\Big[t_{r(i,2)1},s_{r(i,2)1};\dots;t_{r(i,2)\#k_{r(i,2)}},s_{r(i,2)\#k_{r(i,2)}}\Big]}_{\#k_{r(i,2)}},\dots\Big)\Bigg)$$
$$=\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>1\}},\qquad i\in\overline{J}.$$

- Service completion of an order at workstation $m \in \overline{M}$,
  which happens only if there is at least one order,
  i.e. from workstation m, where $\#k_m$ orders are present, an order of type t_{m1} on stage s_{m1} of its route is removed and is sent to the next stage $s_{m1}+1$ of its route, i.e. either

- if $s_{m1} < S(t_{m1})$, it is sent to workstation $r(t_{m1}, s_{m1}+1) \in \overline{M}$, where $\#k_{r(t_{m1},s_{m1}+1)}$ orders have already been present:

$$q\Bigg(\Big(\mathbf{n},\dots,\underbrace{[t_{m1},s_{m1};\dots;t_{m\#k_m},s_{m\#k_m}]}_{\#k_m>0},\dots$$
$$\dots,\underbrace{\Big[t_{r(t_{m1},s_{m1}+1)1},s_{r(t_{m1},s_{m1}+1)1};\dots;t_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}},s_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}}\Big]}_{\#k_{r(t_{m1},s_{m1}+1)}},\dots\Big);$$
$$\Big(\mathbf{n},\dots,\underbrace{[t_{m2},s_{m2};\dots;t_{m\#k_m},s_{m\#k_m}]}_{\#k_m-1},\dots$$
$$\dots,\underbrace{\Big[t_{r(t_{m1},s_{m1}+1)1},s_{r(t_{m1},s_{m1}+1)1};\dots;t_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}},s_{r(t_{m1},s_{m1}+1)\#k_{r(t_{m1},s_{m1}+1)}};t_{m1},s_{m1}+1\Big]}_{\#k_{r(t_{m1},s_{m1}+1)}+1},\dots\Big)\Bigg)$$
$$=\nu_m(\#k_m)\cdot 1_{\{\#k_m>0\}}\cdot 1_{\{s_{m1}<S(t_{m1})\}},\qquad m\in\overline{M},$$

- or if $s_{m1} = S(t_{m1})$, it is sent to the inventory at location $t_{m_1} \in \overline{J}$, where $\#k_{t_{m1}} = 0$ items of raw material have already been present:

$$q\Bigg(\Big(\mathbf{n},\dots,\underbrace{[0]}_{\#k_{t_{m1}}},\dots,\underbrace{[t_{m1},s_{m1};\dots;t_{m\#k_m},s_{m\#k_m}]}_{\#k_m>0},\dots\Big);$$
$$\Big(\mathbf{n},\dots,\underbrace{[t_{m1},1;\dots;t_{m1},1;t_{m1},1]}_{S_{t_{m1}}},\dots,\underbrace{[t_{m2},s_{m2};\dots;t_{m\#k_m},s_{m\#k_m}]}_{\#k_m-1},\dots\Big)\Bigg)$$
$$=\nu_m(\#k_m)\cdot 1_{\{\#k_m>0\}}\cdot 1_{\{s_{m1}=S(t_{m1})\}},\qquad m\in\overline{M}.$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

Proposition 7.3.1. *The strictly positive measure* $\mathbf{x} := \left(x\left(\mathbf{n},\mathbf{k}\right) : \left(\mathbf{n},\mathbf{k}\right) \in E\right)$ *with*

$$x\left(\mathbf{n},\mathbf{k}\right) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}\left(\mathbf{k}\right), \tag{7.3.3}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{7.3.4}$$

$$\widetilde{\theta}(\mathbf{k}) = \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{1_{\{k_j>0\}}} \cdot \prod_{m\in\overline{M}} \prod_{\ell=1}^{\#k_m} \frac{1}{\nu_m(\ell)}, \quad \mathbf{k} \in K, \tag{7.3.5}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z.

Proof. We recall the notation for the inventory-replenishment subsystem. It will sometimes be convenient to use the elaborate notation:

$$\mathbf{k} = \Big(\overbrace{k_1,\dots,k_J}^{\substack{\text{inventories}\\ \text{at locations}}}, \overbrace{k_{J+1},\dots,k_{J+M}}^{\text{supplier network}}\Big)$$

$$= \Bigg(\overbrace{\underbrace{[1,1;\dots;1,1]}_{\#k_1 \text{ items}},\dots,\underbrace{[J,1;\dots;J,1]}_{\#k_J \text{ items}}}^{\substack{\text{inventories}\\ \text{at locations}}},$$

$$\overbrace{\underbrace{\left[t_{(J+1)1}, s_{(J+1)1};\dots;t_{(J+1)\#k_{J+1}}, s_{(J+1)\#k_{J+1}}\right]}_{\#k_{J+1} \text{ orders}},\dots \quad\quad\quad}^{\text{supplier network}}$$

$$\dots,\underbrace{\left[t_{(J+M)1}, s_{(J+M)1};\dots;t_{(J+M)\#k_{J+M}}, s_{(J+M)\#k_{J+M}}\right]}_{\#k_{J+M} \text{ orders}}\Bigg).$$

Note the redundancy of some indicator functions in the global balance equations. We prefer to carry all indicator functions with because it makes it much easier to follow the proof of the stationary distribution.

The global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are given for $(\mathbf{n}, \mathbf{k}) \in E$ from (7.3.1) by

flux out of the state $(\mathbf{n}, \mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if the inventory at this location is not empty (i.e. $\#k_i > 0$) because of the lost sales rule,
- a service completion of a customer at location $i \in \overline{J}$
  if there is at least one customer (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$),
- a completion of an order at workstation $\ell \in \overline{M}$ of the supplier network
  if there is at least one order at this workstation (i.e. $\#k_\ell > 0$):

$$x\Big(\mathbf{n}, \mathbf{k}\Big) \cdot \Big(\underline{\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{\#k_i>0\}}} + \underline{\sum_{i\in\overline{J}} \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}} + \underline{\sum_{\ell\in\overline{M}} \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell>0\}}}\Big)$$

$=$ flux into the state $(\mathbf{n}, \mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one customer at location i (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$):

$$\sum_{i\in\overline{J}} x\Big(\mathbf{n}-\mathbf{e}_i, \mathbf{k}\Big) \cdot \lambda_i \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}$$

- a service completion of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one item and at most $S_i - 1$ items of raw material present at location i (i.e. $0 < \#k_i < S_i$)
  (i.e. a customer departs from location i
  and an item is removed from the associated local inventory there):

$$+\sum_{i\in\overline{J}} x\Big(\mathbf{n}+\mathbf{e}_i, \ldots, \underbrace{[i, 1; i, 1; \ldots; i, 1]}_{\#k_i+1>1}, \ldots\Big) \cdot \mu_i(n_i+1) \cdot 1_{\{\#k_i>0\}} \cdot 1_{\{\#k_i<S_i\}}$$

- a service completion of a customer at location $t_{\ell\#k_\ell} = r(t_{\ell\#k_\ell}, 1) \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one order at workstation ℓ (i.e. $\#k_\ell > 0$) and the order at the tail of the queue at workstation ℓ is in stage 2 of its route (i.e. $s_{\ell\#k_\ell} = 2$)
  (i.e. a customer departs from location $t_{\ell\#k_\ell}$
  and an item is removed from the associated local inventory there,
  and an order is sent to workstation $r(t_{\ell\#k_\ell}, 2) = \ell$)
  (note that $\{s_{\ell\#k_\ell} = 2\}$ implies $\{\#k_{t_{\ell\#k_\ell}} = 0\}$ to hold):

$$+\sum_{\ell\in\overline{M}} x\Big(\mathbf{n}+\mathbf{e}_{t_{\ell\#k_\ell}}, \ldots, \underbrace{[t_{\ell\#k_\ell}, 1]}_{\#k_{t_{\ell\#k_\ell}}+1=1}, \ldots, \underbrace{[t_{\ell 1}, s_{\ell 1}; \ldots; t_{\ell(\#k_\ell-1)}, s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1}, \ldots\Big)$$
$$\cdot \mu_{t_{\ell\#k_\ell}}(n_{t_{\ell\#k_\ell}}+1) \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\{s_{\ell\#k_\ell}=2\}}$$

- a transition of an order of type $t_{\ell\#k_\ell}$ from workstation $r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)$
  to the next workstation of the supplier network
  if in state $(\mathbf{n}, \mathbf{k})$ there is at least one order at workstation ℓ (i.e. $\#k_\ell > 0$) and the order in the tail of the queue at workstation ℓ is not in stage 2 of its route (i.e. $s_{\ell\#k_\ell} > 2$)
  (i.e. an order of type $t_{\ell\#k_\ell}$ is removed from workstation $r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)$
  and is sent to workstation $\ell = r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell})$):

$$+\sum_{\ell\in\overline{M}} x\Big(\mathbf{n},\ldots,\underbrace{\Big[t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1; t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1}, s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\ldots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}, s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\ldots$$
$$\ldots,\underbrace{\big[t_{\ell 1}, s_{\ell 1};\ldots;t_{\ell(\#k_\ell-1)}, s_{\ell(\#k_\ell-1)}\big]}_{\#k_\ell-1},\ldots\Big)$$
$$\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}$$

- a replenishment of the inventory at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there are S_i items of raw material present at location i (i.e. $\#k_i = S_i$)
  (i.e. an order of type i is removed from workstation $r(i, S(i))$
  and is sent to the inventory at location i):

$$+\sum_{i\in\overline{J}} x\Big(\mathbf{n},\ldots,\underbrace{[0]}_{\#k_i-S_i=0},\ldots$$
$$\ldots,\underbrace{\Big[i, S(i); t_{r(i,S(i))1}, s_{r(i,S(i))1};\ldots;t_{r(i,S(i))\#k_{r(i,S(i))}}, s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\ldots\Big)$$
$$\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i=S_i\}}.$$

It has to be shown that the stationary measure (7.3.3) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

Substitution of (7.3.3) and (7.3.4) into the global balance equations directly leads to

$$\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})$$
$$\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}+\sum_{\ell\in\overline{M}}\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\Big)$$

$$=\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}\setminus\{i\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_i(n_i-1)\cdot\widetilde{\theta}(\mathbf{k})\cdot\lambda_i\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}$$

$$+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}\setminus\{i\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_i(n_i+1)\cdot\widetilde{\theta}\Big(\dots,\underbrace{[i,1;i,1;\dots;i,1]}_{\#k_i+1>1},\dots\Big)$$
$$\cdot\,\mu_i(n_i+1)\cdot 1_{\{\#k_i>0\}}\cdot 1_{\{\#k_i<S_i\}}$$

$$+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}\setminus\{i\}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\xi}_i(n_i+1)$$
$$\cdot\,\widetilde{\theta}\Big(\dots,\underbrace{[t_{\ell\#k_\ell},1]}_{\#k_{t_{\ell\#k_\ell}}+1=1},\dots,\underbrace{[t_{\ell1},s_{\ell1};\dots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\dots\Big)$$
$$\cdot\,\mu_{t_{\ell\#k_\ell}}(n_{t_{\ell\#k_\ell}}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}=2\}}$$

$$+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$$
$$\cdot\,\widetilde{\theta}\Big(\dots,\underbrace{\Big[t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\dots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\dots$$
$$\dots,\underbrace{[t_{\ell1},s_{\ell1};\dots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\dots\Big)$$
$$\cdot\,\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\{s_{\ell\#k_\ell}>2\}}$$

$$+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\dots,\underbrace{[0]}_{\#k_i-S_i=0},\dots$$
$$\dots,\underbrace{\Big[i,S(i);t_{r(i,S(i))1},s_{r(i,S(i))1};\dots;t_{r(i,S(i))\#k_{r(i,S(i))}},s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\dots\Big)$$
$$\cdot\,\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i=S_i\}}.$$

Substitution of (7.3.4) leads to

$$
\begin{aligned}
&\left(\prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\\
&\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}+\sum_{\ell\in\overline{M}}\nu_\ell(\#k_\ell)\cdot 1_{\{\#k_\ell>0\}}\Big)\\
=&\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[i,1;i,1;\ldots;i,1]}_{\#k_i+1>1},\ldots\Big)\cdot\lambda_i\cdot 1_{\{\#k_i>0\}}\cdot 1_{\{\#k_i<S_i\}}\\
&+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[t_{\ell\#k_\ell},1]}_{\#k_{t_{\ell\#k_\ell}}+1=1},\ldots,\underbrace{[t_{\ell1},s_{\ell1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\ldots\Big)\\
&\quad\cdot\lambda_{t_{\ell\#k_\ell}}\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\left\{s_{\ell\#k_\ell}=2\right\}}\\
&+\sum_{\ell\in\overline{M}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\Big[t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)1};\ldots\\
&\qquad\underbrace{\ldots;t_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}},s_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1},\ldots\\
&\qquad\ldots,\underbrace{[t_{\ell1},s_{\ell1};\ldots;t_{\ell(\#k_\ell-1)},s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1},\ldots\Big)\\
&\quad\cdot\nu_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell},s_{\ell\#k_\ell}-1)}+1)\cdot 1_{\{\#k_\ell>0\}}\cdot 1_{\left\{s_{\ell\#k_\ell}>2\right\}}\\
&+\sum_{i\in\overline{J}}\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\\
&\quad\cdot\widetilde{\theta}\Big(\ldots,\underbrace{[0]}_{\#k_i-S_i=0},\ldots\\
&\qquad\ldots,\underbrace{\Big[i,S(i);t_{r(i,S(i))1},s_{r(i,S(i))1};\ldots;t_{r(i,S(i))\#k_{r(i,S(i))}},s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1},\ldots\Big)\\
&\quad\cdot\nu_{r(i,S(i))}(\#k_{r(i,S(i))}+1)\cdot 1_{\{\#k_i=S_i\}}.
\end{aligned}
$$

Cancelling $\left(\prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j)\right)$ and the sums with the terms $\mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}$ on both sides of the equation leads to

$$
\begin{aligned}
&\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{\#k_i>0\}} + \sum_{\ell\in\overline{M}} \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell>0\}} \Big) \\
&= \sum_{i\in\overline{J}} \widetilde{\theta}\Big(\ldots, \underbrace{[i,1;i,1;\ldots;i,1]}_{\#k_i+1>1}, \ldots \Big) \cdot \lambda_i \cdot 1_{\{\#k_i>0\}} \cdot 1_{\{\#k_i<S_i\}} \\
&\quad + \sum_{\ell\in\overline{M}} \widetilde{\theta}\Big(\ldots, \underbrace{[t_{\ell\#k_\ell}, 1]}_{\#k_{t_{\ell\#k_\ell}}+1=1}, \ldots, \underbrace{[t_{\ell 1}, s_{\ell 1}; \ldots; t_{\ell(\#k_\ell-1)}, s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1}, \ldots \Big) \\
&\quad \cdot \lambda_{t_{\ell\#k_\ell}} \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\{s_{\ell\#k_\ell}=2\}} \\
&\quad + \sum_{\ell\in\overline{M}} \widetilde{\theta}\Big(\ldots, \Big[t_{\ell\#k_\ell}, s_{\ell\#k_\ell} - 1; t_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)1}, s_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)1}; \ldots \\
&\qquad \underbrace{\ldots; t_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)}}, s_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)\#k_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)}}\Big]}_{\#k_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)}+1}, \ldots \\
&\qquad \ldots, \underbrace{[t_{\ell 1}, s_{\ell 1}; \ldots; t_{\ell(\#k_\ell-1)}, s_{\ell(\#k_\ell-1)}]}_{\#k_\ell-1}, \ldots \Big) \\
&\quad \cdot \nu_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)} + 1) \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\{s_{\ell\#k_\ell}>2\}} \\
&\quad + \sum_{i\in\overline{J}} \widetilde{\theta}\Big(\ldots, \underbrace{[0]}_{\#k_i-S_i=0}, \ldots \\
&\qquad \ldots, \underbrace{\Big[i, S(i); t_{r(i,S(i))1}, s_{r(i,S(i))1}; \ldots, t_{r(i,S(i))\#k_{r(i,S(i))}}, s_{r(i,S(i)-1)\#k_{r(i,S(i))}}\Big]}_{\#k_{r(i,S(i))}+1}, \ldots \Big) \\
&\quad \cdot \nu_{r(i,S(i))}(\#k_{r(i,S(i))} + 1) \cdot 1_{\{\#k_i=S_i\}}.
\end{aligned} \tag{7.3.6}
$$

Substitution of (7.3.5) leads to

$$
\begin{aligned}
&\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{\#k_i>0\}} + \sum_{\ell\in\overline{M}} \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell>0\}} \Big) \\
&= \sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{\#k_i>0\}} \cdot 1_{\{\#k_i<S_i\}} \\
&\quad + \sum_{\ell\in\overline{M}} \widetilde{\theta}(\mathbf{k}) \cdot \nu_\ell(\#k_\ell) \cdot \frac{1}{\lambda_{t_{\ell\#k_\ell}}} \cdot \lambda_{t_{\ell\#k_\ell}} \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\{s_{\ell\#k_\ell}=2\}} \\
&\quad + \sum_{\ell\in\overline{M}} \widetilde{\theta}(\mathbf{k}) \cdot \nu_\ell(\#k_\ell) \cdot \frac{1}{\nu_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)} + 1)} \\
&\qquad \cdot \nu_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)}(\#k_{r(t_{\ell\#k_\ell}, s_{\ell\#k_\ell}-1)} + 1) \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\{s_{\ell\#k_\ell}>2\}} \\
&\quad + \sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot \frac{1}{\nu_{r(i,S(i))}(\#k_{r(i,S(i))} + 1)} \cdot \nu_{r(i,S(i))}(\#k_{r(i,S(i))} + 1) \cdot 1_{\{\#k_i=S_i\}}.
\end{aligned}
$$

$1_{\{\#k_i=S_i\}} = 1$ implies $1_{\{\#k_i>0\}} = 1$. Therefore,

$$\begin{aligned}
&\widetilde{\theta}(\mathbf{k}) \cdot \Big(\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{\#k_i>0\}} \cdot \left(1_{\{\#k_i<S_i\}} + 1_{\{\#k_i=S_i\}}\right) \\
&\qquad + \sum_{\ell\in\overline{M}} \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell>0\}} \cdot \left(1_{\left\{s_{\ell\#k_\ell}=2\right\}} + 1_{\left\{s_{\ell\#k_\ell}>2\right\}}\right)\Big) \\
&= \sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{\#k_i>0\}} \cdot 1_{\{\#k_i<S_i\}} + \sum_{\ell\in\overline{M}} \widetilde{\theta}(\mathbf{k}) \cdot \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\left\{s_{\ell\#k_\ell}=2\right\}} \\
&\quad + \sum_{\ell\in\overline{M}} \widetilde{\theta}(\mathbf{k}) \cdot \nu_\ell(\#k_\ell) \cdot 1_{\{\#k_\ell>0\}} \cdot 1_{\left\{s_{\ell\#k_\ell}>2\right\}} + \sum_{i\in\overline{J}} \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{\#k_i>0\}} \cdot 1_{\{\#k_i=S_i\}}.
\end{aligned}$$

The right-hand side of the last equation is obviously the left-hand side.

Inspection of the system (7.3.6) reveals that it is a "generator equation", i.e. the global balance equation $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$ for a suitably defined ergodic Markov process on state space K with "reduced generator" $\mathbf{Q}_{red}$. Because the Markov process generated by $\mathbf{Q}_{red}$ is irreducible the solution of (7.3.6) is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. □

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

Definition 7.3.2. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := \left(\pi\left(\mathbf{n},\mathbf{k}\right) : (\mathbf{n},\mathbf{k}) \in E\right), \quad \pi\left(\mathbf{n},\mathbf{k}\right) := \lim_{t\to\infty} P\left(Z(t) = (\mathbf{n},\mathbf{k})\right)$$

and the appropriate marginal distributions

$$\xi := \left(\xi\left(\mathbf{n}\right) : \mathbf{n} \in \mathbb{N}_0^J\right), \quad \xi\left(\mathbf{n}\right) := \lim_{t\to\infty} P\left(\left(X_1(t),\ldots,X_J(t)\right) = \mathbf{n}\right),$$

$$\theta := \left(\theta\left(\mathbf{k}\right) : \mathbf{k} \in K\right), \quad \theta\left(\mathbf{k}\right) := \lim_{t\to\infty} P\left(\left(Y_1(t),\ldots,Y_J(t),W_{J+1}(t),\ldots,W_{J+M}(t)\right) = \mathbf{k}\right).$$

Theorem 7.3.3. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi\left(\mathbf{n},\mathbf{k}\right) = \xi(\mathbf{n}) \cdot \theta\left(\mathbf{k}\right), \tag{7.3.7}$$

with

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{7.3.8}$$

$$\theta(\mathbf{k}) = C_\theta^{-1} \cdot \prod_{j\in\overline{J}} \left(\frac{1}{\lambda_j}\right)^{1_{\{k_j>0\}}} \cdot \prod_{m\in\overline{M}} \prod_{\ell=1}^{\#k_m} \frac{1}{\nu_m(\ell)}, \quad \mathbf{k} \in K, \tag{7.3.9}$$

and normalisation constants

$$C_j = \sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)},$$

$$C_\theta = \sum_{\mathbf{k} \in K} \prod_{j \in \overline{J}} \left(\frac{1}{\lambda_j}\right)^{1_{\{k_j > 0\}}} \cdot \prod_{m \in \overline{M}} \prod_{\ell=1}^{\# k_m} \frac{1}{\nu_m(\ell)}.$$

Proof. Z is ergodic, if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ from Proposition 7.3.1 can be normalised (i.e. $\sum_{\mathbf{n} \in \mathbb{N}_0} \sum_{\mathbf{k} \in K} x(\mathbf{n}, \mathbf{k}) < \infty$). Because of Proposition 7.3.1 it holds

$$\begin{aligned}\sum_{\mathbf{n} \in \mathbb{N}_0} \sum_{\mathbf{k} \in K} x(\mathbf{n}, \mathbf{k}) &= \sum_{\mathbf{n} \in \mathbb{N}_0} \widetilde{\xi}(\mathbf{n}) \cdot \sum_{\mathbf{k} \in K} \widetilde{\theta}(\mathbf{k}) \\ &= \left(\prod_{j \in \overline{J}} \sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}\right) \cdot \left(\sum_{\mathbf{k} \in K} \prod_{j \in \overline{J}} \left(\frac{1}{\lambda_j}\right)^{\# k_j} \cdot \prod_{m \in \overline{M}} \prod_{\ell=1}^{\# k_m} \frac{1}{\nu_m(\ell)}\right).\end{aligned}$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 7.3.1 can be normalised if and only if $\sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty$ for all $j \in \overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n}, \mathbf{k}) = \frac{x(\mathbf{n}, \mathbf{k})}{\sum_{\mathbf{n} \in \mathbb{N}_0} \sum_{\mathbf{k} \in K} x(\mathbf{n}, \mathbf{k})},$$

where $x(\mathbf{n}, \mathbf{k})$ is given in Proposition 7.3.1. □

Remark 7.3.4. The expression (7.3.7) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

The explicit formula (7.3.9) for θ shows that in fact there exists a three-term product structure, and that moreover the equilibrium for the integrated model is stratified. In the upper stratum we have two independent vectors for production and inventory-replenishment, the latter splits into two products, a factor for the subsystem comprising the inventories and a factor for the replenishment subsystem.

In the lower stratum each of the three factors of the upper stratum is decomposed completely in "single-component" factors concerning the production servers, the inventories, and the replenishment servers. It should be noted that the factors for the inventories and the replenishment servers do not indicate internal independence, but they are of product form as the celebrated conditionally independent coordinates in the equilibrium of Gordon-Newell networks (see Theorem A.2.6).

Representation (7.3.8) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(\mathbf{n}) = \prod_{j \in \overline{J}} \xi_j(n_j) = \prod_{j \in \overline{J}} C_j^{-1} \cdot \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. In our production network, however, servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (7.3.7) has been unexpected to us.

Comparing our production-inventory-replenishment system with the "Jackson network in a random environment" in [KDO16, Section 4] it turns out that we can interpret the inventory-replenishment subsystem, which contributes via θ to Theorem 7.3.3, as a "random environment" for the production network of nodes $\overline{J}$, which is in this view a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1] we conclude from the hindsight that decoupling of the queueing process $(X_1, \ldots, X_J)$ and the process $(Y_1, \ldots, Y_J, W_{J+1}, \ldots, W_{J+M})$, i.e. the formula (7.3.7), is a consequence of that Theorem 4.1. It should be noted that the three-term structure of the upper stratum of the product form steady state π in Theorem 7.3.3 can not be obtained from the general theory.

Our direct proof of Theorem 7.3.3 is much shorter than embedding the present model into the general framework of [KDO16].

The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

7.3.2. Cost analysis

We consider the following cost structure for inventory, production, and replenishment. The total costs at location $j \in \overline{J}$ consist of shortage costs $c_{ls,j}$ for each customer that is lost, waiting costs $c_{w,j}$ per unit of time for each customer in the system (waiting or in service), capacity costs $c_{s,j}$ per unit of time for providing inventory storage space (e.g. rent, insurance), holding costs $c_{h,j}$ per unit of time for each unit that is kept on inventory. The unit holding costs per item at workstation $m \in \overline{M}$ of the supplier network are $c_{h,m}$. We assume that all of these costs per unit of time are positive.

Therefore, the cost function per unit of time in the respective states is

$$f_{\mathbf{S}}: \ \mathbb{N}_0^J \times K \longrightarrow \mathbb{R}_0^+,$$

$$f_{\mathbf{S}}(\mathbf{n}, \mathbf{k}) = \left(\sum_{j \in \bar{J}} f_{S_j}(n_j, k_j) + \sum_{m \in \overline{M}} f_m(k_m) \right)$$

with the cost functions $f_{S_j}: \ \mathbb{N}_0 \times K_j \longrightarrow \mathbb{R}_0^+$ at location j of the local system state (n_j, k_j) with order-up-to level S_j per unit of time

$$f_{S_j}(n_j, k_j) = c_{w,j} \cdot n_j + c_{s,j} \cdot S_j + c_{h,j} \cdot \#k_j + c_{ls,j} \cdot 1_{\{\#k_j = 0\}}$$

and the cost function $f_m: \ \mathbb{N}_0 \longrightarrow \mathbb{R}_0^+$ at workstation m of the supplier network per unit of time

$$f_m(k_m) = c_{h,m} \cdot \#k_m.$$

We are interested in the long time average costs of the system as a function of the order-up-to levels $\mathbf{S} := \big(S_j : j \in \overline{J}\big)$, which are considered as the main decision variables. The asymptotic average costs for an ergodic system can be calculated as

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f_{\mathbf{S}}(Z(\omega, t))dt = \sum_{(\mathbf{n}, \mathbf{k})} f_{\mathbf{S}}(\mathbf{n}, \mathbf{k}) \cdot \pi(\mathbf{n}, \mathbf{k}) =: \overline{f}(\mathbf{S}) \qquad P - a.s.$$

A cost analysis can be performed as for the basic model in Section 5.4 on page 137.

7.4. $(1, S_j)$-policy with two locations and one workstation

In this section, we study the supply chain as described in Section 7.2 with two locations ($J = 2$, $\overline{J} = \{1, 2\}$) and one workstation ($J + 1 = 3$) at the supplier network ($\overline{M} = \{3\}$). Service times at the supplier are exponentially distributed with parameter $\nu > 0$ and are independent of the queue length. For the control of the inventories we use the $(1, S_j)$-policy. Therefore, if the inventory level at location $j \in \overline{J}$ falls down to the reorder level $r_j = 1$, a replenishment order is instantaneously placed. With each replenishment the local inventory level is restocked to exactly $S_j < \infty$. We assume that $S_j > 0$ and that there is at most one outstanding order ($S_j > 1$). We denote $\mathbf{S} := \left(S_j : j \in \overline{J}\right)$. The supply chain of interest is depicted in Figure 7.4.1.

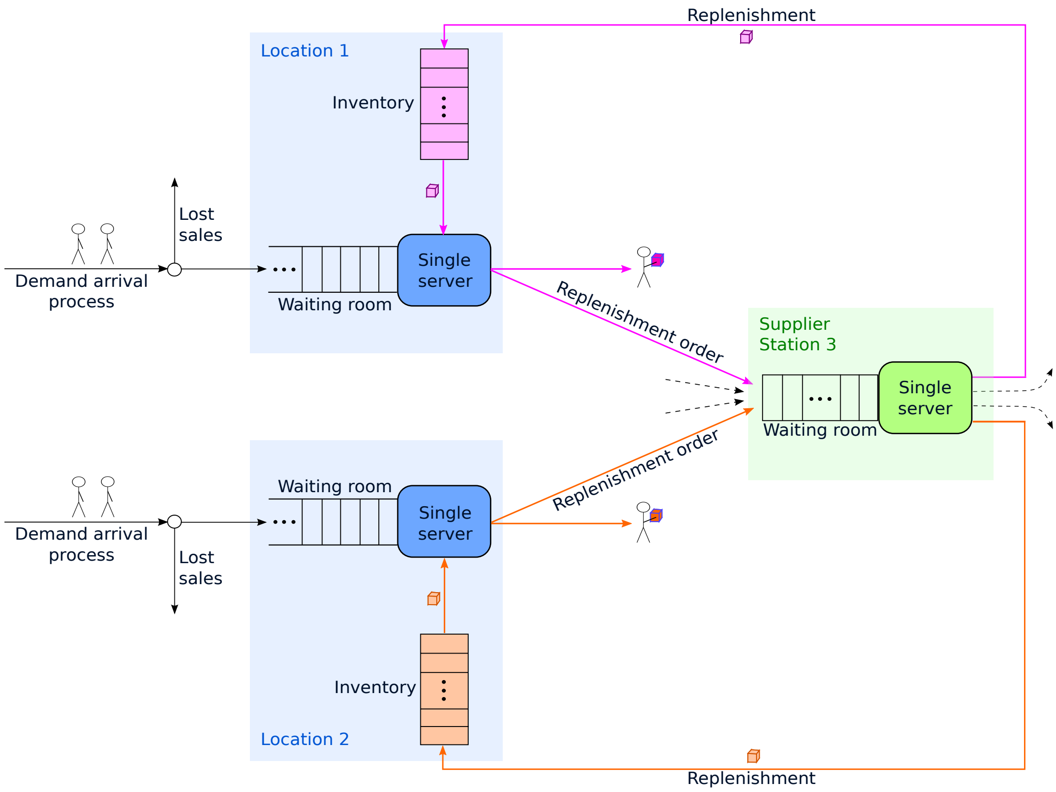


Figure 7.4.1.: Supply chain with two locations, one workstation at the supplier and $(1, S_j)$-policy at location j, $j \in \overline{J}$

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_j(t)$ the number of customers present at location $j \in \overline{J}$ at time $t \geq 0$ either waiting or in service (queue length). By $Y_j(t)$ we denote the contents of the inventory at location $j \in \overline{J}$ at time $t \geq 0$. By $W_{J+1}(t) = W_3(t)$ we denote the sequence of orders at the supplier at time $t \geq 0$.

We denote by K_3 the set of possible states at the supplier (local state space). The state

$$k_3 := \underbrace{[t_1, \dots, t_{\#k_3}]}_{\#k_3 \text{ orders}} \in K_3$$

indicates that there are $\#k_3$ orders at the supplier, on position $p \in \{1, \dots, \#k_3\}$ resides an order of type $t_p \in \overline{J}$. More precisely, t_1 is the order at the head of the line, which is in service, and $t_{\#k_3}$ is the order at the tail of the line. We do not need the stage of its route in the supplier network because the supplier only consists of one workstation.

$$\mathbf{k} = (\#k_1, \#k_2, k_3) \in K \subseteq \prod_{j=1}^{3} K_j,$$

where K_j denotes the local state space at j, $j \in \overline{J} \cup \overline{M}$, and K denotes the feasible states composed of feasible local states.

For $\#k_3 = 0$ we read

$$k_3 = [t_1, \dots, t_{\#k_3}] =: [0].$$

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), X_2(t), Y_1(t), Y_2(t), W_3(t)) \,:\, t \geq 0)\,.$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process, which we assume to be irreducible and regular. The state space of Z is

$$E = \left\{(\mathbf{n}, \mathbf{k}) : \mathbf{n} \in \mathbb{N}_0^J,\, \mathbf{k} \in K\right\}.$$

7.4.1. Limiting and stationary behaviour

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z; \tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(\mathbf{n}, \mathbf{k}) \in E$, where a typical state is

$$(\mathbf{n}, \mathbf{k}) = \Big(n_1, n_2, \#k_1, \#k_2, \underbrace{[t_1, \dots, t_{\#k_3}]}_{=k_3}\Big), \tag{7.4.1}$$

and we will impose necessary restrictions if needed:

- Arrival of a customer at location $i \in \overline{J}$,
  which happens only if the inventory at this location is not empty because of the lost sales rule:
  $$q\left((\mathbf{n}, \mathbf{k}); (\mathbf{n}+\mathbf{e}_i, \mathbf{k})\right) = \lambda_i \cdot 1_{\{\#k_i > 0\}}, \qquad i \in \overline{J}.$$

- Service completion of a customer at location $i \in \overline{J}$,
  which happens only if there is at least one customer at location i and either

  - if there are $\#k_i = 1$ items of raw material present at the inventory at location i, i.e. from location i, where $\#k_i = 1$ items are present, a customer departs and an item of raw material is removed from the associated local inventory and no replenishment order is sent to the supplier because there is at most one outstanding order:

- ⋆ location 1

$$q\Bigg(\Big(\mathbf{n}, \underbrace{\#k_1}_{=1}, \#k_2, [t_1, \dots, t_{\#k_3}]\Big); \Big(\mathbf{n}-\mathbf{e}_1, \underbrace{\#k_1-1}_{=0}, \#k_2, [t_1, \dots, t_{\#k_3}]\Big)\Bigg)$$
$$= \mu_1(n_1) \cdot 1_{\{n_1>0\}} \cdot 1_{\{\#k_1=1\}},$$

- ⋆ location 2

$$q\Bigg(\Big(\mathbf{n}, \#k_1, \underbrace{\#k_2}_{=1}, [t_1, \dots, t_{\#k_3}]\Big); \Big(\mathbf{n}-\mathbf{e}_2, \#k_1, \underbrace{\#k_2-1}_{=0}, [t_1, \dots, t_{\#k_3}]\Big)\Bigg)$$
$$= \mu_2(n_2) \cdot 1_{\{n_2>0\}} \cdot 1_{\{\#k_2=1\}},$$

- if there are $\#k_i = 2$ items of raw material present at the inventory at location i, i.e. from location i, where $\#k_i = 2$ items are present, a customer departs and an item of raw material is removed from the associated local inventory,
in addition a replenishment order of type i is sent to the supplier, where $\#k_3$ have already been present:

  - ⋆ location 1

$$q\Bigg(\Big(\mathbf{n}, \underbrace{\#k_1}_{=2}, \#k_2, [t_1, \dots, t_{\#k_3}]\Big); \Big(\mathbf{n}-\mathbf{e}_1, \underbrace{\#k_1-1}_{=1}, \#k_2, [t_1, \dots, t_{\#k_3}, 1]\Big)\Bigg)$$
$$= \mu_1(n_1) \cdot 1_{\{n_1>0\}} \cdot 1_{\{\#k_1=2\}},$$

  - ⋆ location 2

$$q\Bigg(\Big(\mathbf{n}, \#k_1, \underbrace{\#k_2}_{=2}, [t_1, \dots, t_{\#k_3}]\Big); \Big(\mathbf{n}-\mathbf{e}_2, \#k_1, \underbrace{\#k_2-1}_{=1}, [t_1, \dots, t_{\#k_3}, 2]\Big)\Bigg)$$
$$= \mu_2(n_2) \cdot 1_{\{n_2>0\}} \cdot 1_{\{\#k_2=2\}},$$

- or if $\#k_i > 2$ items of raw material are present at location i,
i.e. from location i, where $\#k_i > 2$ items are present, a customer departs and an item of raw material is removed from the associated local inventory:

  - ⋆ location 1

$$q\Bigg(\Big(\mathbf{n}, \underbrace{\#k_1}_{>2}, \#k_2, [t_1, \dots, t_{\#k_3}]\Big); \Big(\mathbf{n}-\mathbf{e}_1, \underbrace{\#k_1-1}_{>1}, \#k_2, [t_1, \dots, t_{\#k_3}]\Big)\Bigg)$$
$$= \mu_1(n_1) \cdot 1_{\{n_1>0\}} \cdot 1_{\{\#k_1>2\}},$$

  - ⋆ location 2

$$q\Bigg(\Big(\mathbf{n}, \#k_1, \underbrace{\#k_2}_{>2}, [t_1, \dots, t_{\#k_3}]\Big); \Big(\mathbf{n}-\mathbf{e}_2, \#k_1, \underbrace{\#k_2-1}_{>1}, [t_1, \dots, t_{\#k_3}]\Big)\Bigg)$$
$$= \mu_2(n_2) \cdot 1_{\{n_2>0\}} \cdot 1_{\{\#k_2>2\}}.$$

- SERVICE COMPLETION OF AN ORDER AT THE SUPPLIER,
which happens only if there is at least one order (i.e. $\#k_3 > 0$),
i.e. from the workstation at the supplier, where $\#k_3$ orders are present, an order of type t_1 is removed and is sent to the inventory at location $t_1 \in \overline{J}$, where $\#k_{t_1} \le 1$ items of raw material have already been present:

  - location 1

$$q\Bigg(\Big(\mathbf{n}, \underbrace{\#k_1}_{\le 1}, \#k_2, \underbrace{[1, t_2, \dots, t_{\#k_3}]}_{\#k_3>0}\Big); \Big(\mathbf{n}, \underbrace{S_1}_{\#k_1}, \#k_2, \underbrace{[t_2, \dots, t_{\#k_3}]}_{\#k_3-1}\Big)\Bigg)$$
$$= \nu \cdot 1_{\{\#k_3>0\}},$$

  - location 2

$$q\Bigg(\Big(\mathbf{n}, \#k_1, \underbrace{\#k_2}_{\le 1}, \underbrace{[2, t_2, \dots, t_{\#k_3}]}_{\#k_3>0}\Big); \Big(\mathbf{n}, \#k_1, \underbrace{S_2}_{\#k_2}, \underbrace{[t_2, \dots, t_{\#k_3}]}_{\#k_3-1}\Big)\Bigg)$$
$$= \nu \cdot 1_{\{\#k_3>0\}}.$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\,(z;z) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\,(z;\tilde{z}) \qquad \forall z \in E.$$

Proposition 7.4.1. *The strictly positive measure* $\mathbf{x} := (x\,(\mathbf{n},\mathbf{k}) : (\mathbf{n},\mathbf{k}) \in E)$ *with*

$$x\,(\mathbf{n},\mathbf{k}) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}\,(\mathbf{k})\,, \tag{7.4.2}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j\in\overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{7.4.3}$$

and

$$\widetilde{\theta}(\#k_1, \#k_2, [0]) = \frac{\nu\cdot(\lambda_1+\nu)\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+\nu)}{\lambda_1\cdot\lambda_2}, \qquad 1 < \#k_1 \le S_1,\; 1 < \#k_2 \le S_2, \tag{7.4.4}$$

$$\widetilde{\theta}(\#k_1, 1, [2]) = \frac{\nu\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+\nu)}{\lambda_1}, \qquad 1 < \#k_1 \le S_1, \tag{7.4.5}$$

$$\widetilde{\theta}(1, \#k_2, [1]) = \frac{\nu\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+\nu)}{\lambda_2}, \qquad 1 < \#k_2 \le S_2, \tag{7.4.6}$$

$$\widetilde{\theta}(\#k_1, 0, [2]) = \frac{\lambda_2\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+\nu)}{\lambda_1}, \qquad 1 < \#k_1 \le S_1, \tag{7.4.7}$$

$$\widetilde{\theta}(0, \#k_2, [1]) = \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2}, \qquad 1 < \#k_2 \leq S_2, \tag{7.4.8}$$

$$\widetilde{\theta}(1, 1, [1, 2]) = \nu \cdot (\lambda_2 + \nu), \tag{7.4.9}$$

$$\widetilde{\theta}(1, 1, [2, 1]) = \nu \cdot (\lambda_1 + \nu), \tag{7.4.10}$$

$$\widetilde{\theta}(1, 0, [2, 1]) = \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu), \tag{7.4.11}$$

$$\widetilde{\theta}(0, 1, [1, 2]) = \lambda_1 \cdot (\lambda_1 + \lambda_2 + 2\,\nu), \tag{7.4.12}$$

$$\widetilde{\theta}(1, 0, [1, 2]) = \frac{\lambda_2 \cdot \nu \cdot (\lambda_2 + \nu)}{(\lambda_1 + \nu)}, \tag{7.4.13}$$

$$\widetilde{\theta}(0, 1, [2, 1]) = \frac{\lambda_1 \cdot \nu \cdot (\lambda_1 + \nu)}{(\lambda_2 + \nu)}, \tag{7.4.14}$$

$$\widetilde{\theta}(0, 0, [2, 1]) = \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + (\lambda_1 + \nu) \cdot \nu)}{\nu \cdot (\lambda_2 + \nu)}, \tag{7.4.15}$$

$$\widetilde{\theta}(0, 0, [1, 2]) = \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + (\lambda_2 + \nu) \cdot \nu)}{\nu \cdot (\lambda_1 + \nu)} \tag{7.4.16}$$

solves the global balance equations $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ *and is therefore stationary for* Z*.*

Proof. Before proving the proposition we recall the notation for the inventory-replenishment subsystem: It will sometimes be convenient to use the elaborate notation

$$\mathbf{k} = \Big(\overbrace{\#k_1, \#k_2,}^{\substack{\text{inventories}\\ \text{at locations}}} \overbrace{k_3}^{\text{supplier}}\Big) = \Bigg(\overbrace{\#k_1, \#k_2,}^{\substack{\text{inventories}\\ \text{at locations}}} \overbrace{\underbrace{\big[t_1, \dots, t_{\#k_3}\big]}_{\#k_3 \text{ orders}}}^{\text{supplier}}\Bigg).$$

Note the redundancy of some indicator functions in the global balance equations. We prefer to carry all indicator functions with because it makes it much easier to follow the proof of the stationary distribution.

The global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ of the stochastic queueing-inventory process Z are given for $(\mathbf{n},\mathbf{k}) \in E$ from (7.4.1) by

flux out of the state $(\mathbf{n},\mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if the inventory at this location is not empty (i.e. $\#k_i > 0$) because of the lost sales rule,
- a service completion of a customer at location $i \in \overline{J}$
  if there is at least one customer (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$),
- a completion of an order at the supplier
  if there is at least one order (i.e. $\#k_3 > 0$):

$$x\Big(\mathbf{n},\mathbf{k}\Big) \cdot \Big(\underline{\sum_{i\in\overline{J}} \lambda_i \cdot 1_{\{\#k_i>0\}}} + \underline{\sum_{i\in\overline{J}} \mu_i(n_i) \cdot 1_{\{n_i>0\}} \cdot 1_{\{\#k_i>0\}}} + \underline{\nu \cdot 1_{\{\#k_3>0\}}}\Big)$$

= flux into the state $(\mathbf{n},\mathbf{k})$ through:

- an arrival of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n},\mathbf{k})$ there is at least one customer at location i (i.e. $n_i > 0$)
  and the inventory at this location is not empty (i.e. $\#k_i > 0$):
  - location 1

$$x\Big(\mathbf{n}-\mathbf{e}_1,\mathbf{k}\Big) \cdot \lambda_1 \cdot 1_{\{n_1>0\}} \cdot 1_{\{\#k_1>0\}}$$

  - location 2

$$+\,x\Big(\mathbf{n}-\mathbf{e}_2,\mathbf{k}\Big) \cdot \lambda_2 \cdot 1_{\{n_2>0\}} \cdot 1_{\{\#k_2>0\}}$$

- a service completion of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n},\mathbf{k})$ there are at least 2 units
  and at most $S_i - 1$ items of raw material present at location i (i.e. $1 < \#k_i < S_i$)
  (i.e. a customer departs from location i
  and an item is removed from the associated local inventory there
  and no repelenishment order is sent to the supplier because of the $(1, S_j)$-policy):
  - location 1

$$+\,x\Big(\mathbf{n}+\mathbf{e}_1, \#k_1+1, \#k_2, [0]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{1<\#k_1<S_1\}} \cdot 1_{\{\#k_2>1\}} \cdot 1_{\{k_3=[0]\}}$$

$$+\,x\Big(\mathbf{n}+\mathbf{e}_1, \#k_1+1, 1, [2]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{1<\#k_1<S_1\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[2]\}}$$

$$+\,x\Big(\mathbf{n}+\mathbf{e}_1, \#k_1+1, 0, [2]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{1<\#k_1<S_1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2]\}}$$

  - location 2

$$+\,x\Big(\mathbf{n}+\mathbf{e}_2, \#k_1, \#k_2+1, [0]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1>1\}} \cdot 1_{\{1<\#k_2<S_2\}} \cdot 1_{\{k_3=[0]\}}$$

$$+\,x\Big(\mathbf{n}+\mathbf{e}_2, 1, \#k_2+1, [1]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{1<\#k_2<S_2\}} \cdot 1_{\{k_3=[1]\}}$$

$$+\,x\Big(\mathbf{n}+\mathbf{e}_2, 0, \#k_2+1, [1]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{1<\#k_2<S_2\}} \cdot 1_{\{k_3=[1]\}}$$

- a service completion of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is 1 item of raw material present at location i
  (i.e. a customer departs from location i
  and an item is removed from the associated local inventory there
  and a replenishment order of type i is sent to the supplier):

◦ location 1

$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 2, \#k_2, [0]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2>1\}} \cdot 1_{\{k_3=[1]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 2, 1, [2]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[2,1]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 2, 0, [2]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2,1]\}}$$

◦ location 2

$$+ x\Big(\mathbf{n}+\mathbf{e}_2, \#k_1, 2, [0]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1>1\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_2, 1, 2, [1]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[1,2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_2, 0, 2, [1]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[1,2]\}}$$

- a service completion of a customer at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there is 0 item of raw material present at location i
  (i.e. a customer departs from location i
  and an item is removed from the associated local inventory there
  and no replenishment order is sent to the supplier
  because there is at most one outstanding order):

◦ location 1

$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 1, \#k_2, [1]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2>1\}} \cdot 1_{\{k_3=[1]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 1, 1, [1,2]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[1,2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 1, 1, [2,1]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[2,1]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 1, 0, [1,2]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[1,2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_1, 1, 0, [2,1]\Big) \cdot \mu_1(n_1+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2,1]\}}$$

◦ location 2

$$+ x\Big(\mathbf{n}+\mathbf{e}_2, \#k_1, 1, [2]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1>1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_2, 1, 1, [1,2]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[1,2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_2, 1, 1, [2,1]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2,1]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_2, 0, 1, [1,2]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[1,2]\}}$$
$$+ x\Big(\mathbf{n}+\mathbf{e}_2, 0, 1, [2,1]\Big) \cdot \mu_2(n_2+1) \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2,1]\}}$$

- a replenishment of the inventory at location $i \in \overline{J}$
  if in state $(\mathbf{n}, \mathbf{k})$ there are S_i items of raw material present at location i (i.e. $\#k_i = S_i$)
  (i.e. an order of type i is removed from the supplier
  and is sent to the inventory at location i):

∘ location 1

$$+\, x\Big(\mathbf{n}, 0, \#k_2, \underbrace{[1]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=S_1\}} \cdot 1_{\{\#k_2>1\}} \cdot 1_{\{k_3=[0]\}}$$

$$+\, x\Big(\mathbf{n}, 0, 1, \underbrace{[1,2]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=S_1\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[2]\}}$$

$$+\, x\Big(\mathbf{n}, 0, 0, \underbrace{[1,2]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=S_1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2]\}}$$

$$+\, x\Big(\mathbf{n}, 1, \#k_2, \underbrace{[1]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=S_1\}} \cdot 1_{\{\#k_2>1\}} \cdot 1_{\{k_3=[0]\}}$$

$$+\, x\Big(\mathbf{n}, 1, 1, \underbrace{[1,2]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=S_1\}} \cdot 1_{\{\#k_2=1\}} \cdot 1_{\{k_3=[2]\}}$$

$$+\, x\Big(\mathbf{n}, 1, 0, \underbrace{[1,2]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=S_1\}} \cdot 1_{\{\#k_2=0\}} \cdot 1_{\{k_3=[2]\}}$$

∘ location 2

$$+\, x\Big(\mathbf{n}, \#k_1, 0, \underbrace{[2]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1>1\}} \cdot 1_{\{\#k_2=S_2\}} \cdot 1_{\{k_3=[0]\}}$$

$$+\, x\Big(\mathbf{n}, 1, 0, \underbrace{[2,1]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=S_2\}} \cdot 1_{\{k_3=[1]\}}$$

$$+\, x\Big(\mathbf{n}, 0, 0, \underbrace{[2,1]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=S_2\}} \cdot 1_{\{k_3=[1]\}}$$

$$+\, x\Big(\mathbf{n}, \#k_1, 1, \underbrace{[2]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1>1\}} \cdot 1_{\{\#k_2=S_2\}} \cdot 1_{\{k_3=[0]\}}$$

$$+\, x\Big(\mathbf{n}, 1, 1, \underbrace{[2,1]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=1\}} \cdot 1_{\{\#k_2=S_2\}} \cdot 1_{\{k_3=[1]\}}$$

$$+\, x\Big(\mathbf{n}, 0, 1, \underbrace{[2,1]}_{\#k_3+1}\Big) \cdot \nu \cdot 1_{\{\#k_1=0\}} \cdot 1_{\{\#k_2=S_2\}} \cdot 1_{\{k_3=[1]\}}.$$

It has to be shown that the stationary measure (7.4.2) with (7.4.3) and (7.4.4)-(7.4.16) satisfies these global balance equations. Some of the changes are highlighted for reasons of clarity and comprehensibility.

Substitution of (7.4.52) into the global balance equations directly leads to

$$\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\sum_{i\in\overline{J}}\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}+\nu\cdot 1_{\{\#k_3>0\}}\Big)$$

$$=\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\mu_1(n_1)\cdot 1_{\{n_1>0\}}\cdot 1_{\{\#k_1>0\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}(\mathbf{k})\cdot\mu_2(n_2)\cdot 1_{\{n_2>0\}}\cdot 1_{\{\#k_2>0\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1+1,\#k_2,[0]\Big)\cdot\lambda_1\cdot 1_{\{1<\#k_1<S_1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[0]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1+1,1,[2]\Big)\cdot\lambda_1\cdot 1_{\{1<\#k_1<S_1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1+1,0,[2]\Big)\cdot\lambda_1\cdot 1_{\{1<\#k_1<S_1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1,\#k_2+1,[0]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{1<\#k_2<S_2\}}\cdot 1_{\{k_3=[0]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,\#k_2+1,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{1<\#k_2<S_2\}}\cdot 1_{\{k_3=[1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,\#k_2+1,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{1<\#k_2<S_2\}}\cdot 1_{\{k_3=[1]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(2,\#k_2,[0]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(2,1,[2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2,1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(2,0,[2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1,2,[0]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,2,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[1,2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,2,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[1,2]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,\#k_2,[1]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,1,[1,2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[1,2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,1,[2,1]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2,1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,0,[1,2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[1,2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,0,[2,1]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1,1,[2]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,1,[1,2]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[1,2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,1,[2,1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,1,[1,2]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[1,2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,1,[2,1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,\#k_2,[1]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[0]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,1,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,0,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,\#k_2,[1]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[0]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,1,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}}$$
$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,0,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1,0,[2]\Big)\cdot\nu\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[0]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,0,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,0,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(\#k_1,1,[2]\Big)\cdot\nu\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[0]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(1,1,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}}$$

$$+\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)\cdot\widetilde{\theta}\Big(0,1,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}}$$

Cancelling $\left(\prod_{j\in\overline{J}}\widetilde{\xi}_j(n_j)\right)$ and the sums with the terms $\mu_i(n_i)\cdot 1_{\{n_i>0\}}\cdot 1_{\{\#k_i>0\}}$ on both sides of the equation leads to

$$\widetilde{\theta}(\mathbf{k})\cdot\Big(\sum_{i\in\overline{J}}\lambda_i\cdot 1_{\{\#k_i>0\}}+\nu\cdot 1_{\{\#k_3>0\}}\Big) \tag{7.4.17}$$

$$=\widetilde{\theta}\Big(\#k_1+1,\#k_2,[0]\Big)\cdot\lambda_1\cdot 1_{\{1<\#k_1<S_1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[0]\}} \tag{7.4.18}$$

$$+\widetilde{\theta}\Big(\#k_1+1,1,[2]\Big)\cdot\lambda_1\cdot 1_{\{1<\#k_1<S_1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.19}$$

$$+\widetilde{\theta}\Big(\#k_1+1,0,[2]\Big)\cdot\lambda_1\cdot 1_{\{1<\#k_1<S_1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.20}$$

$$+\widetilde{\theta}\Big(\#k_1,\#k_2+1,[0]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{1<\#k_2<S_2\}}\cdot 1_{\{k_3=[0]\}} \tag{7.4.21}$$

$$+\widetilde{\theta}\Big(1,\#k_2+1,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{1<\#k_2<S_2\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.22}$$

$$+\widetilde{\theta}\Big(0,\#k_2+1,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{1<\#k_2<S_2\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.23}$$

$$+\widetilde{\theta}\Big(2,\#k_2,[0]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.24}$$

$$+\widetilde{\theta}\Big(2,1,[2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2,1]\}} \tag{7.4.25}$$

$$+\widetilde{\theta}\Big(2,0,[2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}} \tag{7.4.26}$$

$$+\widetilde{\theta}\Big(\#k_1,2,[0]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.27}$$

$$+\widetilde{\theta}\Big(1,2,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[1,2]\}} \tag{7.4.28}$$

$$+\widetilde{\theta}\Big(0,2,[1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[1,2]\}} \tag{7.4.29}$$

$$+\widetilde{\theta}\Big(1,\#k_2,[1]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.30}$$
$$+\widetilde{\theta}\Big(1,1,[1,2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[1,2]\}} \tag{7.4.31}$$
$$+\widetilde{\theta}\Big(1,1,[2,1]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2,1]\}} \tag{7.4.32}$$
$$+\widetilde{\theta}\Big(1,0,[1,2]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[1,2]\}} \tag{7.4.33}$$
$$+\widetilde{\theta}\Big(1,0,[2,1]\Big)\cdot\lambda_1\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}} \tag{7.4.34}$$

$$+\widetilde{\theta}\Big(\#k_1,1,[2]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.35}$$
$$+\widetilde{\theta}\Big(1,1,[1,2]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[1,2]\}} \tag{7.4.36}$$
$$+\widetilde{\theta}\Big(1,1,[2,1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}} \tag{7.4.37}$$
$$+\widetilde{\theta}\Big(0,1,[1,2]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[1,2]\}} \tag{7.4.38}$$
$$+\widetilde{\theta}\Big(0,1,[2,1]\Big)\cdot\lambda_2\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2,1]\}} \tag{7.4.39}$$

$$+\widetilde{\theta}\Big(0,\#k_2,[1]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[0]\}} \tag{7.4.40}$$
$$+\widetilde{\theta}\Big(0,1,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.41}$$
$$+\widetilde{\theta}\Big(0,0,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.42}$$
$$+\widetilde{\theta}\Big(1,\#k_2,[1]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2>1\}}\cdot 1_{\{k_3=[0]\}} \tag{7.4.43}$$
$$+\widetilde{\theta}\Big(1,1,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=1\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.44}$$
$$+\widetilde{\theta}\Big(1,0,[1,2]\Big)\cdot\nu\cdot 1_{\{\#k_1=S_1\}}\cdot 1_{\{\#k_2=0\}}\cdot 1_{\{k_3=[2]\}} \tag{7.4.45}$$

$$+\widetilde{\theta}\Big(\#k_1,0,[2]\Big)\cdot\nu\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[0]\}} \tag{7.4.46}$$
$$+\widetilde{\theta}\Big(1,0,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.47}$$
$$+\widetilde{\theta}\Big(0,0,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.48}$$
$$+\widetilde{\theta}\Big(\#k_1,1,[2]\Big)\cdot\nu\cdot 1_{\{\#k_1>1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[0]\}} \tag{7.4.49}$$
$$+\widetilde{\theta}\Big(1,1,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=1\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}} \tag{7.4.50}$$
$$+\widetilde{\theta}\Big(0,1,[2,1]\Big)\cdot\nu\cdot 1_{\{\#k_1=0\}}\cdot 1_{\{\#k_2=S_2\}}\cdot 1_{\{k_3=[1]\}}. \tag{7.4.51}$$

Now we will show that (7.4.4)-(7.4.16) satisfies the above equation.

For $\#k_1 = 0$ and $\#k_2 = 0$ and $k_3 = [1,2]$, which corresponds to (7.4.33) and (7.4.38), we obtain with (7.4.16), (7.4.13) and(7.4.12)

$$\widetilde{\theta}(0,0,[1,2]) \cdot \nu = \widetilde{\theta}(1,0,[1,2]) \cdot \lambda_1 + \widetilde{\theta}(0,1,[1,2]) \cdot \lambda_2$$

$$\Leftrightarrow \quad \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+2\,\nu)+(\lambda_2+\nu)\cdot\nu)}{\nu\cdot(\lambda_1+\nu)} \cdot \nu$$
$$= \frac{\lambda_2\cdot\nu\cdot(\lambda_2+\nu)}{(\lambda_1+\nu)} \cdot \lambda_1 + \lambda_1\cdot(\lambda_1+\lambda_2+2\,\nu)\cdot\lambda_2$$

$$\Leftrightarrow \quad \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+2\,\nu)}{(\lambda_1+\nu)} + \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_2+\nu)\cdot\nu}{(\lambda_1+\nu)}$$
$$= \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_2+\nu)\cdot\nu}{(\lambda_1+\nu)} + \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+2\,\nu)}{(\lambda_1+\nu)}.$$

For $\#k_1 = 0$ and $\#k_2 = 0$ and $k_3 = [2,1]$, which corresponds to (7.4.34) and (7.4.39), we obtain with (7.4.15), (7.4.11) and (7.4.14)

$$\widetilde{\theta}(0,0,[2,1]) \cdot \nu = \widetilde{\theta}(1,0,[2,1]) \cdot \lambda_1 + \widetilde{\theta}(0,1,[2,1]) \cdot \lambda_2$$

$$\Leftrightarrow \quad \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+2\,\nu)+(\lambda_1+\nu)\cdot\nu)}{\nu\cdot(\lambda_2+\nu)} \cdot \nu$$
$$= \lambda_2\cdot(\lambda_1+\lambda_2+2\,\nu)\cdot\lambda_1 + \frac{\lambda_1\cdot\nu\cdot(\lambda_1+\nu)}{(\lambda_2+\nu)} \cdot \lambda_2$$

$$\Leftrightarrow \quad \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+2\,\nu)}{(\lambda_2+\nu)} + \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_1+\nu)\cdot\nu}{(\lambda_2+\nu)}$$
$$= \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+2\,\nu)}{(\lambda_2+\nu)} + \frac{\lambda_1\cdot\lambda_2\cdot(\lambda_1+\nu)\cdot\nu}{(\lambda_2+\nu)}.$$

For $\#k_1 = 1$ and $\#k_2 = 1$ and $k_3 = [1,2]$, which corresponds to (7.4.28), we obtain with (7.4.9) and (7.4.6)

$$\widetilde{\theta}(1,1,[1,2]) \cdot (\lambda_1+\lambda_2+\nu) = \widetilde{\theta}(1,2,[1]) \cdot \lambda_2$$

$$\Leftrightarrow \quad \nu\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+\nu) = \frac{\nu\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+\nu)}{\lambda_2} \cdot \lambda_2$$
$$= \nu\cdot(\lambda_2+\nu)\cdot(\lambda_1+\lambda_2+\nu)\,.$$

For $\#k_1 = 1$ and $\#k_2 = 1$ and $k_3 = [2,1]$, which corresponds to (7.4.25), we obtain with (7.4.10) and (7.4.5)

$$\widetilde{\theta}(1,1,[2,1]) \cdot (\lambda_1+\lambda_2+\nu) = \widetilde{\theta}(2,1,[2]) \cdot \lambda_1$$

$$\Leftrightarrow \quad \nu\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+\nu) = \frac{\nu\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+\nu)}{\lambda_1} \cdot \lambda_1$$
$$= \nu\cdot(\lambda_1+\nu)\cdot(\lambda_1+\lambda_2+\nu)\,.$$

For $1 < \#k_1 < S_1$ and $1 < \#k_2 < S_2$ and $k_3 = [0]$, which corresponds to (7.4.18) and (7.4.21), we obtain with (7.4.4)

$$\begin{aligned}
&\widetilde{\theta}(\#k_1, \#k_2, [0]) \cdot (\lambda_1 + \lambda_2) \\
&= \widetilde{\theta}(\#k_1 + 1, \#k_2, [0]) \cdot \lambda_1 + \widetilde{\theta}(\#k_1, \#k_2 + 1, [0]) \cdot \lambda_2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_1 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_2 \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2)\,.
\end{aligned}$$

For $\#k_1 = S_1$ and $\#k_2 = S_2$ and $k_3 = [0]$, which corresponds to (7.4.40), (7.4.43), (7.4.46) and (7.4.49), we obtain with (7.4.4), (7.4.8), (7.4.6), (7.4.7) and (7.4.5)

$$\begin{aligned}
&\widetilde{\theta}(S_1, S_2, [0]) \cdot (\lambda_1 + \lambda_2) \\
&= \widetilde{\theta}\Big(0, S_2, [1]\Big) \cdot \nu + \widetilde{\theta}\Big(1, S_2, [1]\Big) \cdot \nu + \widetilde{\theta}\Big(S_1, 0, [2]\Big) \cdot \nu + \widetilde{\theta}\Big(S_1, 1, [2]\Big) \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2) \\
&= \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \nu + \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \nu \\
&\quad + \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \nu + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \nu \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} + \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2).
\end{aligned}$$

For $\#k_1 = 0$ and $\#k_2 = 1$ and $k_3 = [1, 2]$, which corresponds to (7.4.31) and (7.4.29), we obtain with (7.4.12), (7.4.9) and (7.4.8)

$$\widetilde{\theta}(0, 1, [1, 2]) \cdot (\lambda_2 + \nu) = \widetilde{\theta}(1, 1, [1, 2]) \cdot \lambda_1 + \widetilde{\theta}(0, 2, [1]) \cdot \lambda_2$$

$$\begin{aligned}
\Leftrightarrow \quad \lambda_1 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) \cdot (\lambda_2 + \nu) &= \nu \cdot (\lambda_2 + \nu) \cdot \lambda_1 + \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \lambda_2 \\
&= \lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu)\,.
\end{aligned}$$

For $\#k_1 = 0$ and $\#k_2 = 1$ and $k_3 = [2, 1]$, which corresponds to (7.4.32), we obtain with (7.4.14) and (7.4.10)

$$\widetilde{\theta}(0, 1, [2, 1]) \cdot (\lambda_2 + \nu) = \widetilde{\theta}(1, 1, [2, 1]) \cdot \lambda_1$$

$$\Leftrightarrow \quad \frac{\lambda_1 \cdot \nu \cdot (\lambda_1 + \nu)}{(\lambda_2 + \nu)} \cdot (\lambda_2 + \nu) = \nu \cdot (\lambda_1 + \nu) \cdot \lambda_1$$

$$\Leftrightarrow \quad \lambda_1 \cdot \nu \cdot (\lambda_1 + \nu) = \nu \cdot (\lambda_1 + \nu) \cdot \lambda_1.$$

For $\#k_1 = 0$ and $1 < \#k_2 < S_2$ and $k_3 = [1]$, which corresponds to (7.4.23) and (7.4.30), we obtain with (7.4.8) and (7.4.6)

$$\widetilde{\theta}(0, \#k_2, [1]) \cdot (\lambda_2 + \nu) = \widetilde{\theta}(0, \#k_2 + 1, [1]) \cdot \lambda_2 + \widetilde{\theta}\Big(1, \#k_2, [1]\Big) \cdot \lambda_1$$

$$\begin{aligned}
\Leftrightarrow \qquad & \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot (\lambda_2 + \nu) \\
&= \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \lambda_2 + \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \lambda_1 \\
&= \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot (\lambda_2 + \nu).
\end{aligned}$$

For $\#k_1 = 0$ and $\#k_2 = S_2$ and $k_3 = [1]$, which corresponds to (7.4.30), (7.4.48) and (7.4.51), we obtain with (7.4.8), (7.4.6), (7.4.15) and (7.4.14)

$$\widetilde{\theta}(0, S_2, [1]) \cdot (\lambda_2 + \nu) = \widetilde{\theta}(1, S_2, [1]) \cdot \lambda_1 + \widetilde{\theta}\Big(0, 0, [2, 1]\Big) \cdot \nu + \widetilde{\theta}\Big(0, 1, [2, 1]\Big) \cdot \nu$$

$$\begin{aligned}
\Leftrightarrow \qquad & \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot (\lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \lambda_1 + \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + (\lambda_1 + \nu) \cdot \nu)}{\nu \cdot (\lambda_2 + \nu)} \cdot \nu \\
&\quad + \frac{\lambda_1 \cdot \nu \cdot (\lambda_1 + \nu)}{(\lambda_2 + \nu)} \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \qquad & \frac{\lambda_1 \cdot \lambda_2 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \\
&= \frac{\lambda_1 \cdot \lambda_2 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + \lambda_1 \cdot \lambda_2 \cdot \nu \cdot (\lambda_1 + \nu) + \lambda_1 \cdot \nu^2 \cdot (\lambda_1 + \nu)}{(\lambda_2 + \nu)}
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \qquad & \lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\lambda_1 \cdot \lambda_2 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + \lambda_1 \cdot \nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu)}{(\lambda_2 + \nu)} \\
&= \lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + \lambda_1 \cdot \nu \cdot (\lambda_1 + \nu) \\
&= \lambda_1 \cdot (\lambda_2 \cdot (\lambda_1 + \lambda_2 + \nu) + \nu \cdot (\lambda_1 + \lambda_2 + \nu)) \\
&= \lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)\,.
\end{aligned}$$

For $\#k_1 = 1$ and $\#k_2 = 0$ and $k_3 = [1, 2]$, which corresponds to (7.4.36), we obtain with (7.4.13) and (7.4.9)

$$\widetilde{\theta}(1, 0, [1, 2]) \cdot (\lambda_1 + \nu) = \widetilde{\theta}(1, 1, [1, 2]) \cdot \lambda_2$$

$$\Leftrightarrow \qquad \frac{\lambda_2 \cdot \nu \cdot (\lambda_2 + \nu)}{(\lambda_1 + \nu)} \cdot (\lambda_1 + \nu) = \nu \cdot (\lambda_2 + \nu) \cdot \lambda_2$$

$$\Leftrightarrow \qquad \lambda_2 \cdot \nu \cdot (\lambda_2 + \nu) = \nu \cdot (\lambda_2 + \nu) \cdot \lambda_2.$$

For $\#k_1 = 1$ and $\#k_2 = 0$ and $k_3 = [2,1]$, which corresponds to (7.4.37) and (7.4.26), we obtain with (7.4.11), (7.4.10) and (7.4.7)

$$\widetilde{\theta}(1,0,[2,1]) \cdot (\lambda_1 + \nu) = \widetilde{\theta}(1,1,[2,1]) \cdot \lambda_2 + \widetilde{\theta}(2,0,[2]) \cdot \lambda_1$$

$$\begin{aligned}
\Leftrightarrow \qquad \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) \cdot (\lambda_1 + \nu) &= \nu \cdot (\lambda_1 + \nu) \cdot \lambda_2 + \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_1 \\
&= \nu \cdot (\lambda_1 + \nu) \cdot \lambda_2 + \lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) \cdot (\lambda_1 + \nu).
\end{aligned}$$

For $\#k_1 = 1$ and $1 < \#k_2 < S_2$ and $k_3 = [1]$, which corresponds to (7.4.22) and (7.4.24), we obtain with (7.4.6) and (7.4.4)

$$\begin{aligned}
&\widetilde{\theta}(1, \#k_2, [1]) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \widetilde{\theta}(1, \#k_2 + 1, [1]) \cdot \lambda_2 + \widetilde{\theta}\Big(2, \#k_2, [0]\Big) \cdot \lambda_1
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \qquad &\frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \lambda_2 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_1 \\
&= \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \lambda_2 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \\
&= \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot (\lambda_1 + \lambda_2 + \nu).
\end{aligned}$$

For $\#k_1 = 1$ and $\#k_2 = S_2$ and $k_3 = [1]$, which corresponds to (7.4.24), (7.4.47), (7.4.34) and (7.4.50), we obtain with (7.4.6), (7.4.4), (7.4.11) and (7.4.10)

$$\begin{aligned}
&\widetilde{\theta}(1, S_2, [1]) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \widetilde{\theta}(2, S_2, [0]) \cdot \lambda_1 + \widetilde{\theta}\Big(1, 0, [2,1]\Big) \cdot \nu + \widetilde{\theta}\Big(1, 1, [2,1]\Big) \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \qquad &\frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_1 + \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) \cdot \nu + \nu \cdot (\lambda_1 + \nu) \cdot \nu \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} + \lambda_2 \cdot (\lambda_1 + \lambda_2 + \nu) \cdot \nu + \lambda_2 \cdot \nu^2 + \nu^2 \cdot (\lambda_1 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} + \lambda_2 \cdot (\lambda_1 + \lambda_2 + \nu) \cdot \nu + \nu^2 \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} + \nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu) + \lambda_2 \cdot \nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \\
&= \frac{\nu \cdot (\lambda_1 + \lambda_2 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2}.
\end{aligned}$$

For $1 < \#k_1 < S_1$ and $\#k_2 = 0$ and $k_3 = [2]$, which corresponds to (7.4.20) and (7.4.35), we obtain with (7.4.7) and (7.4.5)

$$\begin{aligned}
&\widetilde{\theta}(\#k_1, 0, [2]) \cdot (\lambda_1 + \nu) \\
&= \widetilde{\theta}(\#k_1 + 1, 0, [2]) \cdot \lambda_1 + \widetilde{\theta}\Big(\#k_1, 1, [2]\Big) \cdot \lambda_2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad &\frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot (\lambda_1 + \nu) \\
&= \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_1 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_2 \\
&= \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_1 + \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \nu. \\
&= \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot (\lambda_1 + \nu).
\end{aligned}$$

For $1 < \#k_1 < S_1$ and $\#k_2 = 1$ and $k_3 = [2]$, which corresponds to (7.4.19) and (7.4.27), we obtain with (7.4.5) and (7.4.4)

$$\begin{aligned}
&\widetilde{\theta}(\#k_1, 1, [2]) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \widetilde{\theta}(\#k_1 + 1, 1, [2]) \cdot \lambda_1 + \widetilde{\theta}\Big(\#k_1, 2, [0]\Big) \cdot \lambda_2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad &\frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_1 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_1 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot (\lambda_1 + \lambda_2 + \nu).
\end{aligned}$$

For $1 < \#k_1 < S_1$ and $\#k_2 = S_2$ and $k_3 = [0]$, which corresponds to (7.4.18), (7.4.46) and (7.4.49), we obtain with (7.4.4), (7.4.7) and (7.4.5)

$$\begin{aligned}
&\widetilde{\theta}(\#k_1, S_2, [0]) \cdot (\lambda_1 + \lambda_2) \\
&= \widetilde{\theta}(\#k_1 + 1, S_2, [0]) \cdot \lambda_1 + \widetilde{\theta}\Big(\#k_1, 0, [2]\Big) \cdot \nu + \widetilde{\theta}\Big(\#k_1, 1, [2]\Big) \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad &\frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_1 + \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \nu \\
&\quad + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \nu \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_1 + \frac{(\lambda_2 + \nu) \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \nu \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_1 + \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_2.
\end{aligned}$$

For $\#k_1 = S_1$ and $\#k_2 = 0$ and $k_3 = [2]$, which corresponds to (7.4.35), (7.4.42) and (7.4.45), we obtain with (7.4.7), (7.4.5), (7.4.16) and (7.4.13)

$$\begin{aligned}
&\widetilde{\theta}(S_1, 0, [2]) \cdot (\lambda_1 + \nu) \\
&= \widetilde{\theta}(S_1, 1, [2]) \cdot \lambda_2 + \widetilde{\theta}\Big(0, 0, [1, 2]\Big) \cdot \nu + \widetilde{\theta}\Big(1, 0, [1, 2]\Big) \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot (\lambda_1 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot \lambda_2 + \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + (\lambda_2 + \nu) \cdot \nu)}{\nu \cdot (\lambda_1 + \nu)} \cdot \nu \\
&\quad + \frac{\lambda_2 \cdot \nu \cdot (\lambda_2 + \nu)}{(\lambda_1 + \nu)} \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \\
&= \frac{\lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + \lambda_1 \cdot \lambda_2 \cdot (\lambda_2 + \nu) \cdot \nu + \lambda_2 \cdot \nu^2 \cdot (\lambda_2 + \nu)}{(\lambda_1 + \nu)}
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + \lambda_2 \cdot \nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \nu)}{(\lambda_1 + \nu)} \\
&= \lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + \lambda_2 \cdot \nu \cdot (\lambda_2 + \nu) \\
&= \lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2 + \nu) + \lambda_1 \cdot \lambda_2 \cdot \nu + \lambda_2 \cdot \nu \cdot (\lambda_2 + \nu) \\
&= \lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2 + \nu) + \lambda_2 \cdot \nu \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)\,.
\end{aligned}$$

For $\#k_1 = S_1$ and $\#k_2 = 1$ and $k_3 = [2]$, which corresponds to (7.4.27), (7.4.41) and (7.4.44), we obtain with (7.4.5), (7.4.4), (7.4.12) and (7.4.9)

$$\widetilde{\theta}(S_1, 1, [2]) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(S_1, 2, [0]) \cdot \lambda_2 + \widetilde{\theta}\Big(0, 1, [1, 2]\Big) \cdot \nu + \widetilde{\theta}\Big(1, 1, [1, 2]\Big) \cdot \nu$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_2 + \lambda_1 \cdot (\lambda_1 + \lambda_2 + 2\,\nu) \cdot \nu + \nu \cdot (\lambda_2 + \nu) \cdot \nu \\
&= \frac{(\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} + \lambda_1 \cdot (\lambda_1 + \lambda_2 + \nu) + \lambda_1 \cdot \nu + \nu \cdot (\lambda_2 + \nu) \\
&= \frac{(\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} + \lambda_1 \cdot (\lambda_1 + \lambda_2 + \nu) + \nu \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{(\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1} + (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu) \\
&= \frac{(\lambda_1 + \nu) \cdot ((\lambda_2 + \nu) + \lambda_1) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1}.
\end{aligned}$$

For $\#k_1 = S_1$ and $1 < \#k_2 < S_2$ and $k_3 = [0]$, which corresponds to (7.4.21), (7.4.40) and (7.4.43), we obtain with (7.4.4), (7.4.8) and (7.4.6)

$$\begin{aligned}
&\widetilde{\theta}(S_1, \#k_2, [0]) \cdot (\lambda_1 + \lambda_2) \\
&= \widetilde{\theta}(S_1, \#k_2 + 1, [0]) \cdot \lambda_2 + \widetilde{\theta}\Big(0, \#k_2, [1]\Big) \cdot \nu + \widetilde{\theta}\Big(1, \#k_2, [1]\Big) \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2) \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_2 + \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \nu \\
&\quad + \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot \nu \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot \lambda_2 + \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \\
&= \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2} \cdot (\lambda_1 + \lambda_2).
\end{aligned}$$

□

Recall that the system is irreducible and regular. Therefore, if Z has a stationary and limiting distribution, this is uniquely defined.

For the queueing-inventory process Z on state space E holds Definition 7.3.2 with $J = 2$ and $M = 1$. More precisely, it holds

$$\xi(\mathbf{n}) := \lim_{t\to\infty} P\left((X_1(t), X_2(t)) = \mathbf{n}\right),$$

$$\theta(\mathbf{k}) := \lim_{t\to\infty} P\left((Y_1(t), Y_2(t), W_3(t)) = \mathbf{k}\right).$$

Theorem 7.4.2. *The queueing-inventory process Z is ergodic if and only if for $j \in \overline{J}$*

$$\sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty.$$

If Z is ergodic, then its unique limiting and stationary distribution is

$$\pi(\mathbf{n}, \mathbf{k}) = \xi(\mathbf{n}) \cdot \theta(\mathbf{k}), \tag{7.4.52}$$

with

$$\xi(\mathbf{n}) = \prod_{j \in \overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0,\ j \in \overline{J}, \tag{7.4.53}$$

and

$$\theta(\#k_1, \#k_2, [0]) = C_\theta^{-1} \cdot \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1 \cdot \lambda_2}, \qquad 1 < \#k_1 \le S_1,\ 1 < \#k_2 \le S_2,$$

$$\theta(\#k_1, 1, [2]) = C_\theta^{-1} \cdot \frac{\nu \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1}, \qquad 1 < \#k_1 \le S_1,$$

$$\theta(1, \#k_2, [1]) = C_\theta^{-1} \cdot \frac{\nu \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2}, \qquad 1 < \#k_2 \le S_2,$$

$$\theta(\#k_1, 0, [2]) = C_\theta^{-1} \cdot \frac{\lambda_2 \cdot (\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_1}, \qquad 1 < \#k_1 \le S_1,$$

$$\theta(0, \#k_2, [1]) = C_\theta^{-1} \cdot \frac{\lambda_1 \cdot (\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + \nu)}{\lambda_2}, \qquad 1 < \#k_2 \le S_2,$$

$$\theta(1, 1, [1, 2]) = C_\theta^{-1} \cdot \nu \cdot (\lambda_2 + \nu)\,,$$

$$\theta(1, 1, [2, 1]) = C_\theta^{-1} \cdot \nu \cdot (\lambda_1 + \nu)\,,$$

$$\theta(1, 0, [2, 1]) = C_\theta^{-1} \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2 + 2\,\nu)\,,$$

$$\theta(0, 1, [1, 2]) = C_\theta^{-1} \cdot \lambda_1 \cdot (\lambda_1 + \lambda_2 + 2\,\nu)\,,$$

$$\theta(1, 0, [1, 2]) = C_\theta^{-1} \cdot \frac{\lambda_2 \cdot \nu \cdot (\lambda_2 + \nu)}{(\lambda_1 + \nu)},$$

$$\theta(0, 1, [2, 1]) = C_\theta^{-1} \cdot \frac{\lambda_1 \cdot \nu \cdot (\lambda_1 + \nu)}{(\lambda_2 + \nu)},$$

$$\theta(0, 0, [2, 1]) = C_\theta^{-1} \cdot \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_2 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + (\lambda_1 + \nu) \cdot \nu)}{\nu \cdot (\lambda_2 + \nu)},$$

$$\theta(0, 0, [1, 2]) = C_\theta^{-1} \cdot \frac{\lambda_1 \cdot \lambda_2 \cdot ((\lambda_1 + \nu) \cdot (\lambda_1 + \lambda_2 + 2\,\nu) + (\lambda_2 + \nu) \cdot \nu)}{\nu \cdot (\lambda_1 + \nu)}$$

and normalisation constants

$$C_j = \sum_{n_j \in \mathbb{N}_0} \prod_{\ell=1}^{n_j} \left(\frac{\lambda_j}{\mu_j(\ell)} \right)$$

and C_θ, which can be calculated by the normalisation condition.

Proof. Z is ergodic, if and only if the strictly positive measure $\mathbf{x}$ of the global balance equation $\mathbf{x} \cdot \mathbf{Q} = \mathbf{0}$ from Proposition 7.4.1 can be normalised (i.e. $\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k}) < \infty$). Because of Proposition 7.4.1 it holds

$$\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k}) = \sum_{\mathbf{n}\in\mathbb{N}_0} \widetilde{\xi}(\mathbf{n}) \cdot \sum_{\mathbf{k}\in K} \widetilde{\theta}\,(\mathbf{k})$$
$$= \left(\prod_{j\in\overline{J}} \sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} \right) \cdot \left(\sum_{\mathbf{k}\in K} \widetilde{\theta}\,(\mathbf{k}) \right).$$

Hence, since K is finite, the measure $\mathbf{x}$ from Proposition 7.4.1 can be normalised if and only if $\sum_{n_j\in\mathbb{N}_0} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)} < \infty$ for all $j \in \overline{J}$.

Consequently, if the process is ergodic, the limiting and stationary distribution π is given by

$$\pi(\mathbf{n},\mathbf{k}) = \frac{x(\mathbf{n},\mathbf{k})}{\sum_{\mathbf{n}\in\mathbb{N}_0} \sum_{\mathbf{k}\in K} x(\mathbf{n},\mathbf{k})},$$

where $x(\mathbf{n},\mathbf{k})$ is given in Proposition 7.4.1. □

Remark 7.4.3. The expression (7.4.52) shows that the two-component production-inventory-replenishment system is separable, the steady states of the production network and the inventory-replenishment complex decouple asymptotically.

Representation (7.4.53) shows that the equilibrium for the production subsystem decomposes in true independent coordinates. A product structure of the stationary distribution as

$$\xi(\mathbf{n}) = \prod_{j\in\overline{J}} \xi_j(n_j) = \prod_{j\in\overline{J}} C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}$$

is commonly found for standard Jackson networks (see Theorem A.2.2) and their relatives. In Jackson networks servers are "non-idling", i.e. they are always busy as long as customers are present at the respective node. In our production network, however, servers may be idle while there are customers waiting because a replenishment needs to arrive first. Consequently, the product form (7.4.52) has been unexpected to us.

Comparing our production-inventory-replenishment system with the "Jackson network in a random environment" in [KDO16, Section 4] it turns out that we can interpret the inventory-replenishment subsystem, which contributes via θ to Theorem 7.4.2, as a "random environment" for the production network of nodes $\overline{J}$, which is in this view a Jackson network of parallel servers (for more details see Appendix A.3). Taking into account the results of [KDO16, Theorem 4.1] we conclude from the hindsight that decoupling of the queueing process (X_1, X_2) and the process (Y_1, Y_2, W_3), i.e. the formula (7.4.53), is a consequence of that Theorem 4.1.

Our direct proof of Theorem 7.4.2 is much shorter than embedding the present model into the general framework of [KDO16].

The structural properties from Section 2.6.1 (ergodicity) and Section 2.6.5 (insensitivity and robustness) hold word-by-word for this integrated system as well.

8. Directions for future research

The directions of future research in this chapter are relevant for the models in Part II as well.

Early reservation of items in the inventory

To count the size of inventory, the most common assumption in the literature is to remove one item of raw material from the inventory when a served customer departs from the system. With some suitable simplifications this leads to simple regimes for the production-inventory systems and partly even to explicit steady state analysis.

For example Schwarz and her coauthors ([Sch04], [SSD$^+$06]) investigated $M/M/1$ systems with attached inventory, exponentially distributed lead times and lost sales. They derived the steady state distributions of the joint queue length and inventory process in explicit product form and calculated important measures of system performance. Schwarz and her coauthors noted that their “single threshold” policies will only yield suboptimal results, and can be considered as heuristic decision making with simple policies. However, Berman and Kim ([BK01], [BK04]) proved that an optimal replenishment policy is of “multiple threshold” structure and has to take the inventory size and the queue length into consideration as decision variables. In hindsight, this seems to be intuitive, since when a long queue has been accumulated, it should be a better policy to send out a replenishment order earlier than in the case of a nearly empty queue.

A natural, but still simple modelling approach is the following as shown in Figure 8.0.1: Whenever a customer arrives at the system, one item of raw material is reserved for him. In the system’s state the free inventory, this means the items of raw material which are not reserved, is counted.

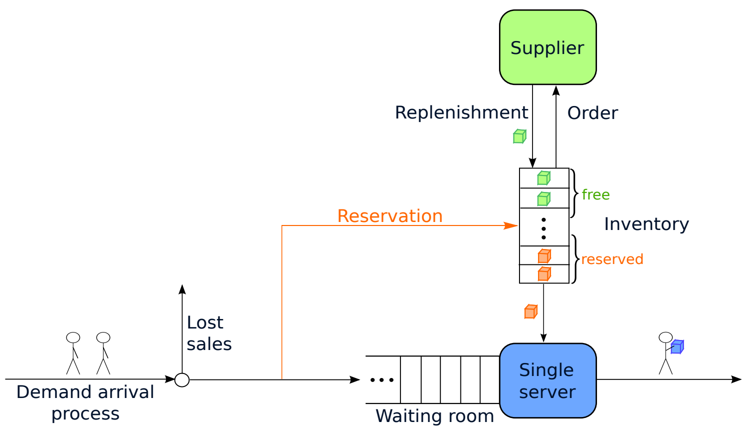


Figure 8.0.1.: Production-inventory system: Early reservation of items

This issue seems to be important since such “order promising policies” are becoming increasingly important in the supply chain management. These policies define the rules for accepting or rejecting the orders in dependence on the availability of the product and the capacity. Slotnick gives a review of the literature about order acceptance and scheduling from a problem-oriented perspective in [Slo11].

Examples for order promising policies are available-to-promise (ATP) and capable-to-promise (CTP).

The ATP concept “(...) directly links available resources, including both material and capacity, with customer orders and, thus, affects the overall performance of a supply chain” [CZB01, p. 477].

The concept of CTP is an extension of the traditional ATP. “In particular, the generalisation consists in the integration of additional information about capacities and intermediate product inventories in multi-stage production systems.” [GK15, p. 136] There exist some extensions of these approaches in the literature. For example Gössinger and Kalkowski present an extension of the capable-to-promise approach in [GK15] and Pibernik presents an advanced available-to-promise scheme (AATP) in [Pib05].

For more information about order promising policies and literature reviews see for example [CZB01, Section 1.2, pp. 478f.] and [GK15, pp. 135ff.].

However, this modelling approach has not been investigated analytically in the literature so far. We started to concentrate on the investigation of these production-inventory models in the master thesis [Ott13, pp. 17-162].

Main contributions in [Ott13, pp. 17-162]*:* Markov process models for production-inventory systems with infinite waiting room under standard inventory control policies and a general randomized reorder scheme. The criterion for ergodicity (stability) is determined and it has been proven that the steady state distribution of the queueing-inventory process is not of product form. Furthermore, the marginal steady state distribution of the free inventory process has been determined and the probability of an empty queue at the production system is found. An iterative matrix scheme to calculate the steady state probabilities has been derived. However, the needed initial value for the scheme is not explicitly at hand yet. If this initial value can be computed, this scheme could be considered as a new computational algorithm to derive steady state distributions for a specific class of matrix analytic models.

Separable approximation of non-separable systems

The models considered in this thesis can be extended in several directions. Unfortunately, for most extensions closed form solutions for the stationary distribution can not be expected. Thus, highly expensive numerical or approximative computations have to be performed.

For example, it would be interesting to analyse the models where the assumption of lost sales is relaxed to partial backordering or backordering. As mentioned by Schwarz “(...) things become much more involved if the lost sales assumption is abandoned allowing now for demand being backordered” [Sch04, p. 87]. Furthermore, van Dijk argued that practical phenomena such as blocking or dynamic routing, breakdown features, capacity constraints and prioritisations violate closed-form expressions (cf. [Dij98, p. 296]).

Further interesting extensions are the following:

- arbitrarily distributed service or lead times,
- interruptions in queue because of e.g. maintenance or repair (for a literature review about queues with interruptions we refer to [KPC12]),
- transfer of the items between the inventories at the locations,
- multiple supply options (for a literature review about multiple-supplier inventory models in supply chain management we refer to [Min03]).

As mentioned before in Section 1.1, an important question is "Can we use our product form results to obtain simple product form bounds for systems with unknown non-product form stationary distributions?". Van Dijk and his coauthors (e.g. [Dij11b, Section 1,7, pp. 62f.], [Dij98, pp. 311ff.], [DK92], [DW89]) show that a product form modification turns out to be quite fruitful to provide product form bounds for the throughput of a unsolvable ($\equiv$ unknown stationary distribution) queueing-inventory system.

Part II.

Production-inventory systems with priority classes

9. Introduction

Up to now, one of the key assumptions of queueing-inventory models in literature is that customers are indistinguishable. In practice, however, customers have different characteristics and/or priorities, which leads to systems where this assumption does not hold. In the following chapters, we therefore consider systems with two customer classes with different priorities. In this chapter, we focus on queue-length-independent arrival and service rates.

9.1. Related literature and own contributions

Wang and his coauthors [WBSW15] state there are three main motivations for prioritisation:

- Different customers may have a different willingness to pay for the same product.
- Customers may require different products or services, where some of these products are more profitable than others.
- Different service levels may substantially affect long-term profitability.

"Modelling the effects of prioritization due to the first and third motivations can be achieved with identical service time distributions for different segments." [WBSW15, p. 733]

According to Isotupa [Iso15, p. 411], there are many situations where it would be financially beneficial for a supplier to provide different levels of service to different customers.

Wang and his coauthors [WBSW15, p. 733] write that customers with different service requests have different priorities in many practical applications. "For example, contact centres prioritize phone calls over emails; renewals of driver's licenses require a photograph and thus typically take longer than renewals of car licenses; at airports, processing times of the aircrew are shorter than those of air travellers, who have a lower priority" [WBSW15, p. 733]. Our models do not have different service requests, but the priorities have an impact on the costs.

According to Yadavalli and his coauthors [YAJ15a], patients with serious illnesses are given higher priority than the other patients opting for routine checks or else in multi-speciality hospitals.

Another example from Liu and his couauthors [LXC13, pp. 1544f.] is that orders with long term contracts have higher priority than unscheduled order since they may bear lower shortage cost than the booked orders.

For a literature review about inventory control systems with multiple customer classes, we refer to Isotupa ([Iso11, Section 2, pp. 3ff.], [Iso15, Section 1, pp. 411ff.]) and Arslan and his coauthors [AGR07, pp. 1486ff.].

It can be differentiated between priority disciplines that regulate customer arrivals and priority disciplines that regulate customer services. In this chapter, we will combine both disciplines.

In a sequence of papers ([Iso06], [Iso07], [Iso11], [IS13], [Iso15]), Isotupa investigates inventory systems with the most commonly used replenishment policies — (r, Q)-policy and base stock policy — with two classes of customers regarding the priority discipline that regulates customer arrivals, or more precisely, the lost sales conditions for ordinary customers. Examples are inventory systems of spare parts in the airline or shipping industries and spare parts for refinery equipment, since the equipment is categorized in many different classes and different service levels are defined for each type (cf. [Iso11, pp. 1f.]).

According to Isotupa ([Iso07], [Iso11], [Iso15]) the systems under investigation can differ in terms of the priority discipline that regulates customer arrivals or more precisely the lost sales conditions for ordinary customers:

Case 1: During the time the inventory is depleted, arriving priority customers and arriving ordinary customers are lost (i.e. no differentiation between customers).

Case 2: During the time the inventory is equal to or less than a threshold level s satisfying $0 < s \leq r$ or $r < s \leq Q$, arriving ordinary customers are lost. If the inventory level is zero, demands due to both types of customers are lost.[1]

Isotupa derives a closed form expression for the stationary distribution of the inventory levels, the performance measures and the long-run expected cost rate. She proves in [Iso07], [Iso11] and [Iso15] that under certain conditions the rationing policy yields lower costs and provides better service levels to both types of customers than the policy where all customers are treated alike.

Liu and his coauthors ([CZL12], [LXC13], [LFW14]) introduce a flexible admission control with a priority parameter $0 \leq p \leq 1$ "for controlling the application of priority" [LFW14, p. 181] for the inventory system of Isotupa with case 2 of the priority discipline. The priority parameter p indicates the probability with which the arrivals of ordinary customers are treated like the arrivals of priority customers. If $p = 1$, there is no priority in regulating arrivals. If $p = 0$, there is a strict priority in regulating arrivals. In the last case, their model is the same as the model of Isotupa [Iso07]. They derive the stationary distribution of the inventory levels and some performance measures. To obtain the optimal inventory control policy they construct a mixed integer optimization problem in [LXC13] and [LFW14] and develop an efficient searching algorithm in [CZL12].

However, the above systems are inventory systems that do not include a production system. In an inventory system, when a customer's demand arrive, "(...) it is typically required to do some processing on the inventory item (e.g. retrieval, preparation, packing, and loading) before delivering it to the customer" [HHS11]. To the best of our knowledge,

[1] Isotupa does not consider the case $Q < s \leq r + Q$ without any reasoning.

production-inventory systems with different classes of customers were first considered by Zhao and Lian [ZL11]. They investigate a system with Poisson arrivals, exponentially distributed service and lead times under backordering. They find a priority service rule to minimise the long-run expected waiting cost by a dynamic programming method. They formulate the model as a level-dependent quasi-birth-and-death process such that the steady state probability distribution of their production-inventory systems can be computed by the Bright-Taylor algorithm.

To the best of our knowledge, this extension has not been considered so far in the context of a production system with attached inventory and lost sales. Therefore, we dedicated parts of our research to production-inventory systems with two classes of customers and inventory management under lost sales where the customers' arrivals are regulated by a flexible admission control.

Jeganathan and his coauthors investigate in a sequence of papers (e.g. [JKA16], [Jeg15], [YAJ15a]) queueing-inventory systems with two classes of customers. They consider models with impatient customers, an optional second service and a mixed priority service (non-preemptive priority and preemptive priority).

Li and Zhao [LZ09] investigate a preemptive priority queueing system (without inventory) with two classes of customers and an exponential single server who serves the two classes of customers at potentially different rates.

Our main contributions are the following:
We develop a Markov process description for a production-inventory system with two classes of customers and inventory management under lost sales with (r, Q)-policy in Chapter 10 and base stock policy in Chapter 11 where customers' arrivals are regulated by a flexible admission control. The global balance equations and the existence of a stationary distribution are investigated. Furthermore, we consider the special case of zero service time, which results in a pure inventory system, and determine the stationary distribution.

Our system with (r, Q)-policy is an extension of the production-inventory system of Zhao and Lian [ZL11] with two classes of customers in the context of lost sales with a priority parameter.

Furthermore, our work with (r, Q)-policy is an extension of the investigations of Isotupa ([Iso06], [Iso07], [Iso11]), Liu and his coauthors ([LFW14], [LXC13], [CZL12]): In their models no production processes are considered. Therefore, their models are special cases of our model when the service time is set to zero. In addition, the priority parameter is not an issue of the investigations in the models of Isotupa. Moreover, Isotupa assumes that the threshold level s is not greater than the order quantity Q without any reasoning and Liu and his authors assume that the threshold level is equal to the reorder level. Hence, we consider a threshold level that can also be greater than the order quantity.

Moreover, our study with base stock policy is an extension to the investigations of Isotupa [Iso15]. In her model is no production system. In terms of our production-inventory scenario, we will arrive at her model when setting the service time to zero. In her model the replenishment rate depends on the number of pending orders. In Remark 11.1.10, we explain how we can extend our model so that the replenishment lead time depends on the number of orders at the supplier.

On the other hand, it is an extension of the production-inventory systems under (r, Q)-policy described by Schwarz and her coauthors [SSD+06]: Their study is limited to one class of customers. Our results extend their setting to two classes of customers with different priorities and flexible admission control.

9.2. Description of the general model

The supply chain of interest is depicted in Figure 9.2.1 and consists of priority and ordinary customers, a production system (a single server with two unlimited waiting rooms), an inventory and a supplier.

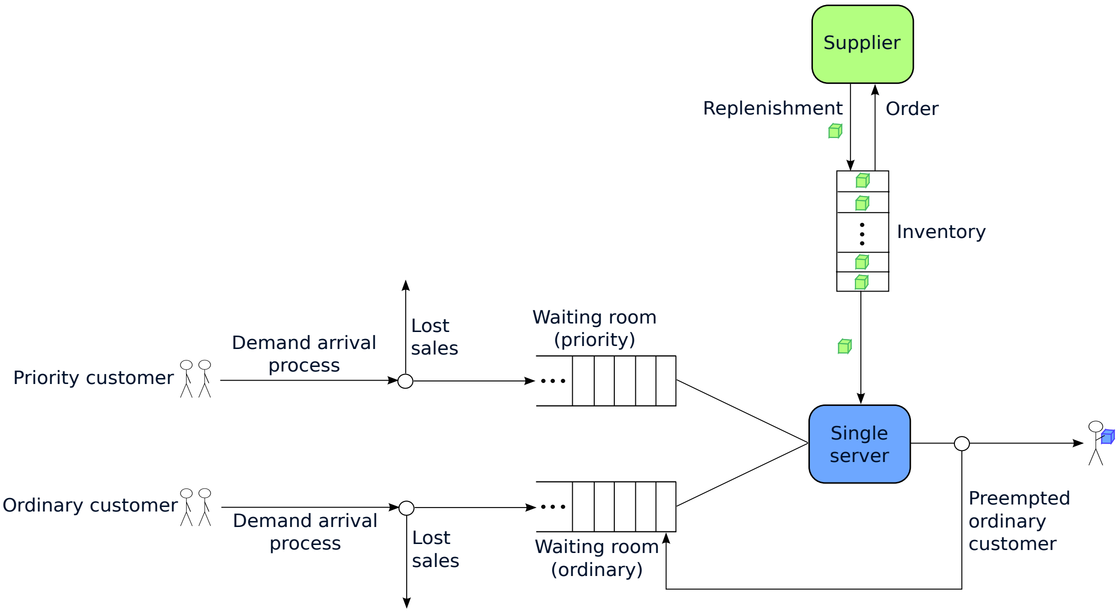


Figure 9.2.1.: The production-inventory system with two customer classes

The production system manufactures according to customers' demand on a make-to-order basis. There are two types of customers — priority customers and ordinary customers. $\overline{C} = \{1, 2\}$ is the set of customer classes where 1 is the type of priority customers and 2 is the type of ordinary customers — the smaller the number, the higher the priority. Priority customers arrive according to a Poisson process with rate $\lambda_1 > 0$ and ordinary customers arrive according to a Poisson process with rate $\lambda_2 > 0$.

Customers' arrivals are regulated by a flexible admission control with priority parameter p, $0 \le p \le 1$: If the inventory is depleted all arriving customers are rejected ("lost sales"). If the on-hand inventory is greater than a prescribed threshold level s, customers of both classes are admitted to enter the system[2]. If the on-hand inventory reaches or falls

[2]The threshold level s is greater than zero and smaller than the maximal size of the inventory. Hence, in Chapter 10, where the (r, S)-policy is considered, for the threshold level holds $0 < s < r + Q$ and in Chapter 11, where the base stock policy is considered, for the threshold level holds $0 < s < b$.

below the threshold level s, priority customers still enter the system but ordinary customers are allowed to enter only with probability p and are rejected with probability $1-p$.

There is a single server with two separate infinite waiting rooms — one waiting room for priority customers (priority queue) and one waiting room for ordinary customers (ordinary queue) both under a FCFS regime. If both customer queues are not empty, the server needs to decide which one of them should be served. The choice is made according to a prescribed priority discipline. An overview of various priority disciplines can be found in [Jai68, p. 53].

In Chapter 10 and Chapter 11, the preemptive resume discipline is considered. According to this discipline a newly arriving priority customer interrupts immediately an ongoing service of an ordinary customer. The preempted ordinary customer returns to the head of his ordinary queue and then has to wait until the priority queue is exhausted before he reenters service. The preempted customer resumes service from the point of interruption so that his service time upon reentry has been reduced by the amount of time the customer has already spent in service (cf. [Mil58, p. 1]). Since it is assumed that the service time is exponential, the ordinary customer requires on its reentry stochastically the same amount of service as it required on its earlier entry. Thus, the preemptive resume discipline is equal to the preemptive repeat-identical discipline where the preempted customer requires the same amount of service on its reentry as he required on his earlier entry (cf. [Jai68, p. 53]).

Each customer needs exactly one item from the inventory for service. The service time for both types of customers is exponentially distributed with parameter $\mu > 0$. If the server is ready to serve a customer, who is at the head of the line, and the inventory is not depleted, the service begins immediately. Otherwise, the service starts at the instant of time when the next replenishment arrives at the inventory.

A served customer departs from the system immediately and the associated item is removed from the inventory at this time instant. It is assumed that the transportation time between the production system and the inventory is negligible.

An outside supplier replenishes raw material to the inventory according to a continuous review replenishment policy. At each decision epoch a replenishment policy determines whether a replenishment order is placed or not, and how many items are ordered. Admissible decision epochs are arrival and departure epochs. The systems under investigation differ with respect to the replenishment policy in the following way:

- (r, Q)-policy (Chapter 10),
- base stock policy (Chapter 11).

We investigate systems where the replenishment lead time is exponentially distributed with parameter $\nu > 0$.

It is assumed that transmission times for orders are negligible and set to zero.

All service times, inter-arrival times and replenishment lead times constitute an independent family of random variables.

10. Production-inventory system with (r, Q)-policy

In this chapter, we study the production-inventory system with two types of arriving customers, a flexible admission control and (r, Q)-policy, as described in Section 9.2.

The inventory is controlled by the (r, Q)-policy. This means, if the on-hand inventory falls down to a prefixed value $r \geq 0$, a replenishment order is placed instantaneously. The size of the order is fixed to $Q < \infty$ units of raw material. We assume that $r < Q$ (this "ensures that there is no perpetual shortage" [LFW14, p. 1545]) and that there is at most one outstanding order. The maximal size of the inventory is $r + Q$.

Remark 10.0.1. With respect to economic aspects, if $\nu = \infty$ a reorder level $r > 0$ under any policy does not make sense, since r items of raw material are never touched by the customers and remain in stock forever (cf. [SSD$^+$06, p. 63]).

Customer arrivals are regulated by a *flexible admission control* with priority parameter p, $0 \leq p \leq 1$: If the inventory is depleted all arriving customers are rejected ("lost sales"). If the on-hand inventory is greater than a prescribed threshold level s, $0 < s < r + Q$, the customers of both classes are admitted to enter the system. If the on-hand inventory reaches or falls below the threshold level s, all priority customers still enter the system, but ordinary customers are allowed to enter only with probability p and are rejected with probability $1 - p$.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_1(t)$ the number of priority customers present in the system at time $t \geq 0$, and by $X_2(t)$ the number of ordinary customers in the system at time $t \geq 0$ either waiting or in service (queue length). Since the customer in service will always be of the priority class when at least one priority customer is present, the value of the vector $(X_1(t), X_2(t))$ determines uniquely the type of the customer in service at time $t \geq 0$, if any. Moreover, by $Y(t)$ we denote the on-hand inventory at time $t \geq 0$.

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), X_2(t), Y(t)) : t \geq 0).$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process. The state space of Z is

$$E = \left\{(n_1, n_2, k) : (n_1, n_2) \in \mathbb{N}_0^2,\ k \in \{0, \ldots, r + Q\}\right\},$$

where $r + Q$ is the maximal size of the inventory, which depends on the replenishment policy.

10.1. Properties of the stationary system

In this section, we assume that the queueing-inventory process Z is ergodic.

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(n_1, n_2, k) \in E$:

$$\begin{aligned}
q((n_1,n_2,k);(n_1+1,n_2,k)) &= \lambda_1 \cdot 1_{\{k>0\}}, \\
q((n_1,n_2,k);(n_1,n_2+1,k)) &= p\,\lambda_2 \cdot 1_{\{0<k\leq s\}} + \lambda_2 \cdot 1_{\{k>s\}}, \\
q((n_1,n_2,k);(n_1-1,n_2,k-1)) &= \mu \cdot 1_{\{n_1>0\}} \cdot 1_{\{k>0\}}, \\
q((n_1,n_2,k);(n_1,n_2-1,k-1)) &= \mu \cdot 1_{\{n_1=0\}} \cdot 1_{\{n_2>0\}} \cdot 1_{\{k>0\}}, \\
q((n_1,n_2,k);(n_1,n_2,k+Q)) &= \nu \cdot 1_{\{k\leq r\}}.
\end{aligned}$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

Definition 10.1.1. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := \left(\pi\left(n_1,n_2,k\right) : \left(n_1,n_2,k\right) \in E\right) \quad \text{by } \pi\left(n_1,n_2,k\right) := \lim_{t\to\infty} P\left(Z(t) = \left(n_1,n_2,k\right)\right).$$

The global balance equations $\pi \cdot \mathbf{Q} = \mathbf{0}$ of the ergodic queueing-inventory process Z are for $(n_1, n_2, k) \in E$ given by

$$\begin{aligned}
&\pi(n_1,n_2,k) \cdot \big((\lambda_1 + p\,\lambda_2) \cdot 1_{\{0<k\leq s\}} + (\lambda_1+\lambda_2) \cdot 1_{\{k>s\}} \\
&\qquad\qquad\qquad\qquad + \mu \cdot 1_{\{n_1+n_2>0\}} \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k\leq r\}}\big) \\
&= \pi(n_1-1,n_2,k) \cdot \lambda_1 \cdot 1_{\{n_1>0\}} \cdot 1_{\{k>0\}} \\
&\quad + \pi(n_1,n_2-1,k) \cdot p\,\lambda_2 \cdot 1_{\{n_2>0\}} \cdot 1_{\{0<k\leq s\}} + \pi(n_1,n_2-1,k) \cdot \lambda_2 \cdot 1_{\{n_2>0\}} \cdot 1_{\{k>s\}} \\
&\quad + \pi(n_1+1,n_2,k+1) \cdot \mu \cdot 1_{\{k<r+Q\}} + \pi(n_1,n_2+1,k+1) \cdot \mu \cdot 1_{\{n_1=0\}} \cdot 1_{\{k<r+Q\}} \\
&\quad + \pi(n_1,n_2,k-Q) \cdot \nu \cdot 1_{\{k\geq Q\}}.
\end{aligned} \tag{10.1.1}$$

Let (X_1, X_2, Y) be a random variable that is distributed according to the queueing-inventory process in equilibrium. Then, Y is a random variable that is distributed according to the inventory process in equilibrium and X_1 resp. X_2 are random variables that are respectively distributed according to the queue length processes of priority resp. ordinary customers in equilibrium.

Proposition 10.1.2. *For the queueing-inventory process Z holds the following equilibrium of probability flows*

$$\underbrace{P(Y > 0) \cdot \lambda_1}_{\substack{\textit{effective arrival rate}\\ \textit{of priority customers}}} = \underbrace{P(X_1 > 0, Y > 0) \cdot \mu}_{\substack{\textit{effective departure rate}\\ \textit{of priority customers}}}, \tag{10.1.2}$$

$$\underbrace{P(0 < Y \leq s) \cdot p\,\lambda_2 + P(Y > s) \cdot \lambda_2}_{\substack{\textit{effective arrival rate}\\ \textit{of ordinary customers}}} = \underbrace{P(X_1 = 0, X_2 > 0, Y > 0) \cdot \mu}_{\substack{\textit{effective departure rate}\\ \textit{of ordinary customers}}} \tag{10.1.3}$$

$$\begin{aligned}&\underbrace{P(Y > 0) \cdot \lambda_1 + P(0 < Y \leq s) \cdot p\,\lambda_2 + P(Y > s) \cdot \lambda_2}_{\substack{\textit{effective arrival rate}\\ \textit{of customers}}}\\ &= \underbrace{P(X_1 + X_2 > 0, Y > 0) \cdot \mu}_{\substack{\textit{effective departure rate}\\ \textit{of customers}}}\end{aligned} \tag{10.1.4}$$

and

$$\begin{aligned}&P(X_1 = n_1, Y > 0) \cdot \lambda_1\\ &= P(X_1 = n_1 + 1, Y > 0) \cdot \mu, \qquad n_1 \in \mathbb{N}_0,\end{aligned} \tag{10.1.5}$$

$$\begin{aligned}&P(X_2 = n_2, 0 < Y \leq s) \cdot p\,\lambda_2 + P(X_2 = n_2, Y > s) \cdot \lambda_2\\ &= P(X_1 = 0, X_2 = n_2 + 1, Y > 0) \cdot \mu, \qquad n_2 \in \mathbb{N}_0,\end{aligned} \tag{10.1.6}$$

$$\begin{aligned}&P(X_1 + X_2 = n, 0 < Y \leq s) \cdot (\lambda_1 + p\,\lambda_2)\\ &+P(X_1 + X_2 = n, Y > s) \cdot (\lambda_1 + \lambda_2)\\ &= P(X_1 + X_2 = n + 1, Y > 0) \cdot \mu, \qquad n \in \mathbb{N}_0.\end{aligned} \tag{10.1.7}$$

Remark 10.1.3. The effective departure rates are usually called throughputs.

Proof. The equations can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a) on page 259.

For $n_1 \in \mathbb{N}_0$, equation (10.1.5) can be proven by a cut, which divides E into complementary sets according to the queue length of priority customers that is less than or equal to n_1 or greater than n_1, i.e. into the sets

$$\Big\{(m_1, m_2, k) : m_1 \in \{0, 1, \dots, n_1\},\ m_2 \in \mathbb{N}_0,\ k \in \{0, \dots, r+Q\}\Big\},$$

$$\Big\{(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) : \widetilde{m}_1 \in \mathbb{N}_0 \setminus \{0, 1, \dots, n_1\},\ \widetilde{m}_2 \in \mathbb{N}_0,\ \widetilde{k} \in \{0, \dots, r+Q\}\Big\}, \quad n_1 \in \mathbb{N}_0.$$

Then, the following holds for $n_1 \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{m_1=0}^{n_1} \sum_{m_2=0}^{\infty} \sum_{k=0}^{r+Q} \sum_{\widetilde{m}_1=n_1+1}^{\infty} \sum_{\widetilde{m}_2=0}^{\infty} \sum_{\widetilde{k}=0}^{r+Q} \pi(m_1, m_2, k) \cdot q((m_1, m_2, k); (\widetilde{m}_1, \widetilde{m}_2, \widetilde{k})) \\
&= \sum_{\widetilde{m}_1=n_1+1}^{\infty} \sum_{\widetilde{m}_2=0}^{\infty} \sum_{\widetilde{k}=0}^{r+Q} \sum_{m_1=0}^{n_1} \sum_{m_2=0}^{\infty} \sum_{k=0}^{r+Q} \pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) \cdot q((\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}); (m_1, m_2, k)) \\
\Leftrightarrow\quad &\underbrace{\sum_{m_1=n_1}^{n_1} \sum_{m_2=0}^{\infty} \sum_{k=1}^{r+Q} \pi(m_1, m_2, k) \cdot \lambda_1}_{=P(X_1=n_1, Y>0)\cdot\lambda_1} = \underbrace{\sum_{\widetilde{m}_1=n_1+1}^{n_1+1} \sum_{\widetilde{m}_2=0}^{\infty} \sum_{\widetilde{k}=1}^{r+Q} \pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) \cdot \mu}_{=P(X_1=n_1+1, Y>0)\cdot\mu}.
\end{aligned}$$

Hence, for $n_1 \in \mathbb{N}_0$ holds (10.1.5)

$$P(X_1 = n_1, Y > 0) \cdot \lambda_1 = P(X_1 = n_1 + 1, Y > 0) \cdot \mu.$$

Summation of the equations (10.1.5) over $n_1 \in \mathbb{N}_0$ leads to

$$\sum_{n_1=0}^{\infty} P(X_1 = n_1, Y > 0) \cdot \lambda_1 = \sum_{n_1=0}^{\infty} P(X_1 = n_1 + 1, Y > 0) \cdot \mu,$$

which is equivalent to (10.1.2)

$$P(Y > 0) \cdot \lambda_1 = P(X_1 > 0, Y > 0) \cdot \mu.$$

For $n_2 \in \mathbb{N}_0$, equation (10.1.6) can be proven by a cut, which divides E into complementary sets according to the queue length of ordinary customers that is less than or equal to n_2 or greater than n_2, i.e. into the sets

$$\Big\{(m_1, m_2, k) : m_1 \in \mathbb{N}_0,\ m_2 \in \{0, 1, \dots, n_2\},\ k \in \{0, \dots, r+Q\}\Big\},$$

$$\Big\{(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) : \widetilde{m}_1 \in \mathbb{N}_0,\ \widetilde{m}_2 \in \mathbb{N}_0 \setminus \{0, 1, \dots, n_2\},\ \widetilde{k} \in \{0, \dots, r+Q\}\Big\}, \quad n_2 \in \mathbb{N}_0.$$

Then, the following holds for $n_2 \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{m_1=0}^{\infty} \sum_{m_2=0}^{n_2} \sum_{k=0}^{r+Q} \sum_{\widetilde{m}_1=0}^{\infty} \sum_{\widetilde{m}_2=n_2+1}^{\infty} \sum_{\widetilde{k}=0}^{r+Q} \pi(m_1, m_2, k) \cdot q((m_1, m_2, k); (\widetilde{m}_1, \widetilde{m}_2, \widetilde{k})) \\
&= \sum_{\widetilde{m}_1=0}^{\infty} \sum_{\widetilde{m}_2=n_2+1}^{\infty} \sum_{\widetilde{k}=0}^{r+Q} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{n_2} \sum_{k=0}^{r+Q} \pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) \cdot q((\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}); (m_1, m_2, k)) \\
\Leftrightarrow \quad & \underbrace{\sum_{m_1=0}^{\infty} \sum_{m_2=n_2}^{n_2} \sum_{k=1}^{s} \pi(m_1, m_2, k) \cdot p\, \lambda_2}_{=P(X_2=n_2, 0<Y\le s)\cdot p\,\lambda_2} + \underbrace{\sum_{m_1=0}^{\infty} \sum_{m_2=n_2}^{n_2} \sum_{k=s+1}^{r+Q} \pi(m_1, m_2, k) \cdot \lambda_2}_{=P(X_2=n_2, Y>s)\cdot\lambda_2} \\
&= \underbrace{\sum_{\widetilde{m}_1=0}^{0} \sum_{\widetilde{m}_2=n_2+1}^{n_2+1} \sum_{\widetilde{k}=1}^{r+Q} \pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) \cdot \mu}_{=P(X_1=0, X_2=n_2+1, Y>0)\cdot\mu} .
\end{aligned}$$

Thus, for $n_2 \in \mathbb{N}_0$ holds (10.1.6)

$$\begin{aligned}
&P(X_2 = n_2, 0 < Y \le s) \cdot p\, \lambda_2 + P(X_2 = n_2, Y > s) \cdot \lambda_2 \\
&= P(X_1 = 0, X_2 = n_2 + 1, Y > 0) \cdot \mu.
\end{aligned}$$

Summation of the equations (10.1.6) over $n_2 \in \mathbb{N}_0$ leads to

$$\begin{aligned}
&\sum_{n_2=0}^{\infty} P(X_2 = n_2, 0 < Y \le s) \cdot p\, \lambda_2 + \sum_{n_2=0}^{\infty} P(X_2 = n_2, Y > s) \cdot \lambda_2 \\
&= \sum_{n_2=0}^{\infty} P(X_1 = 0, X_2 = n_2 + 1, Y > 0) \cdot \mu,
\end{aligned}$$

which is equivalent to (10.1.3)

$$P(0 < Y \le s) \cdot p\, \lambda_2 + P(Y > s) \cdot \lambda_2 = P(X_1 = 0, X_2 > 0, Y > 0) \cdot \mu.$$

For $n \in \mathbb{N}_0$, equation (10.1.7) can be proven by a cut, which divides E into complementary sets according to the total queue length of customers that is less than or equal to n or greater than n, i.e. into the sets

$$\Big\{(m_1,m_2,k) : m_1,m_2 \in \mathbb{N}_0, (m_1+m_2) \in \{0,1,\dots,n\},\ k \in \{0,\dots,r+Q\}\Big\},$$

$$\Big\{(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k}) : \widetilde{m}_1,\widetilde{m}_2 \in \mathbb{N}_0, (\widetilde{m}_1+\widetilde{m}_2) \in \mathbb{N}_0 \setminus \{0,1,\dots,n\},\ \widetilde{k} \in \{0,\dots,r+Q\}\Big\},\ n \in \mathbb{N}_0.$$

Then, the following holds for $n \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{m_1+m_2=0}^{n}\sum_{k=0}^{r+Q}\ \sum_{\widetilde{m}_1+\widetilde{m}_2=n+1}^{\infty}\sum_{\widetilde{k}=0}^{r+Q}\pi(m_1,m_2,k)\cdot q((m_1,m_2,k);(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k}))\\
=&\sum_{\widetilde{m}_1+\widetilde{m}_2=n+1}^{\infty}\sum_{\widetilde{k}=0}^{r+Q}\ \sum_{m_1+m_2=0}^{n}\sum_{k=0}^{r+Q}\pi(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k})\cdot q((\widetilde{m}_1,\widetilde{m}_2,\widetilde{k});(m_1,m_2,k))\\
\Leftrightarrow\quad &\underbrace{\sum_{m_1+m_2=n}^{n}\sum_{k=1}^{s}\pi(m_1,m_2,k)\,(\lambda_1+p\,\lambda_2)}_{=P(X_1+X_2=n,0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)}+\underbrace{\sum_{m_1+m_2=n}^{n}\sum_{k=s+1}^{r+Q}\pi(m_1,m_2,k)\cdot(\lambda_1+\lambda_2)}_{=P(X_1+X_2=n,Y>s)\cdot(\lambda_1+\lambda_2)}\\
=&\underbrace{\sum_{\widetilde{m}_1+\widetilde{m}_2=n+1}^{n+1}\sum_{\widetilde{k}=1}^{r+Q}\pi(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k})\cdot\mu\,.}_{=P(X_1+X_2=n+1,Y>0)\cdot\mu}
\end{aligned}$$

Hence, for $n \in \mathbb{N}_0$ holds (10.1.7)

$$\begin{aligned}
&P(X_1+X_2=n,0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)+P(X_1+X_2=n,Y>s)\cdot(\lambda_1+\lambda_2)\\
=&\,P(X_1+X_2=n+1,Y>0)\cdot\mu.
\end{aligned}$$

Summation of the equations (10.1.7) over $n \in \mathbb{N}_0$ leads to

$$\begin{aligned}
&\sum_{n=0}^{\infty}P(X_1+X_2=n,0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)+\sum_{n=0}^{\infty}P(X_1+X_2=n,Y>s)\cdot(\lambda_1+\lambda_2)\\
=&\sum_{n=0}^{\infty}P(X_1+X_2=n+1,Y>0)\cdot\mu,
\end{aligned}$$

which is equivalent to (10.1.4)

$$\begin{aligned}
&P(0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)+P(Y>s)\cdot(\lambda_1+\lambda_2)\\
=&\,P(X_1+X_2>0,Y>0)\cdot\mu\\
\Leftrightarrow\quad &P(Y>0)\cdot\lambda_1+P(0<Y\le s)\cdot p\,\lambda_2+P(Y>s)\cdot\lambda_2\\
=&\,P(X_1+X_2>0,Y>0)\cdot\mu.
\end{aligned}$$

□

Remark 10.1.4.

(a) From (10.1.5) follows for $n_1 \in \mathbb{N}_0$,

$$P(X_1 = n_1, Y > 0) = P(X_1 = 0, Y > 0) \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1}. \qquad (10.1.8)$$

(b) In an $M/M/c$ queue with two priority classes under a preemptive priority discipline without inventory, the priority customers form a classic $M/M/c$ queue (cf. [WBSW15]). However, our queueing-inventory system with $c = 1$ (one server) is not so easy to solve since two customer classes have to share the same inventory and therefore, the priority customers do not form a classic $M/M/1$ queue.

(c) Rearranging (10.1.4) shows that the probability that the inventory is not depleted is given by

$$\begin{aligned} P(Y > 0) = {} & P(X_1 + X_2 > 0, Y > 0) \cdot \frac{\mu}{\lambda_1} \\ & - P(0 < Y \leq s) \cdot \frac{p\,\lambda_2}{\lambda_1} - P(Y > s) \cdot \frac{\lambda_2}{\lambda_1} \end{aligned}$$

and from (10.1.2) follows

$$P(Y > 0) = P(X_1 > 0, Y > 0) \cdot \frac{\mu}{\lambda_1}.$$

Corollary 10.1.5. *For the conditional distribution of the queue length process of priority customers conditioned on $\{Y > 0\}$ holds for $n_1 \in \mathbb{N}_0$*

$$P(X_1 = n_1 | Y > 0) = P(X_1 = 0 | Y > 0) \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1}$$

with

$$P(X_1 = 0 | Y > 0) = \left(1 - \frac{\lambda_1}{\mu}\right).$$

This shows that the conditional queue length process of priority customers under the condition that the inventory is not empty has in equilibrium the same structure as a birth-and-death process with birth-rates λ_1 and death-rates μ.

Proof. Equation (10.1.8) implies for $n_1 \in \mathbb{N}_0$

$$\begin{aligned} P(X_1 = n_1 | Y > 0) &= \frac{P(X_1 = n_1, Y > 0)}{P(Y > 0)} \\ &\overset{(10.1.8)}{=} \frac{P(X_1 = 0, Y > 0)}{P(Y > 0)} \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1} \\ &= P(X_1 = 0 | Y > 0) \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1} \end{aligned}$$

and the normalizing condition leads to

$$P(X_1 = 0 | Y > 0) = \left(\sum_{n_1=1}^{\infty} \left(\frac{\lambda_1}{\mu}\right)^{n_1}\right)^{-1} = \left(1 - \frac{\lambda_1}{\mu}\right).$$

□

Proposition 10.1.6. *The probability that a replenishment order is outstanding fulfils the following equalities:*

$$P(Y<r+1)=\frac{\mu}{\nu}\cdot P(X_1+X_2>0,Y=Q) \tag{10.1.9}$$
$$=\frac{\mu}{\nu}\cdot P(X_1+X_2>0,Y=r+1) \tag{10.1.10}$$

and

$$P(Y<r+1)=\frac{\mu}{\nu\cdot Q}\cdot P(X_1+X_2>0,Y>0). \tag{10.1.11}$$

Proof. Equation (10.1.9) can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a) on page 259. For $Q>r$, it can be proven by a cut, which divides E into complementary sets according to the size of the inventory that is less than or equal to $Q-1$ or greater than $Q-1$, i.e. into the sets

$$\Big\{(n_1,n_2,k):(n_1,n_2)\in\mathbb{N}_0^2,\ k\in\{0,\dots,Q-1\}\Big\},$$
$$\Big\{(\widetilde{n}_1,\widetilde{n}_2,\widetilde{k}):(\widetilde{n}_1,\widetilde{n}_2)\in\mathbb{N}_0^2,\ \widetilde{k}\in\{Q,\dots,r+Q\}\Big\},\quad r<Q.$$

Then, the following holds for $r<Q$

$$\begin{aligned}
&\sum_{n_1+n_2=0}^{\infty}\sum_{k=0}^{Q-1}\sum_{\widetilde{n}_1+\widetilde{n}_2=0}^{\infty}\sum_{\widetilde{k}=Q}^{r+Q}\pi(n_1,n_2,k)\cdot q((n_1,n_2,k);(\widetilde{n}_1,\widetilde{n}_2,\widetilde{k}))\\
&=\sum_{\widetilde{n}_1+\widetilde{n}_2=0}^{\infty}\sum_{\widetilde{k}=Q}^{r+Q}\sum_{n_2+n_2=0}^{\infty}\sum_{k=0}^{Q-1}\pi(\widetilde{n}_1,\widetilde{n}_2,\widetilde{k})\cdot q((\widetilde{n}_1,\widetilde{n}_2,\widetilde{k});(n_1,n_2,k))\\
\Leftrightarrow\quad&\underbrace{\sum_{n_1+n_2=0}^{\infty}\sum_{k=0}^{r}\pi(n_1,n_2,k)\cdot\nu}_{=P(Y<r+1)\cdot\nu}=\underbrace{\sum_{\widetilde{n}_1+\widetilde{n}_2=1}^{\infty}\sum_{\widetilde{k}=Q}^{Q}\pi(\widetilde{n}_1,\widetilde{n}_2,\widetilde{k})\cdot\mu}_{=P(X_1+X_2>0,Y=Q)\cdot\mu}.
\end{aligned}$$

Therefore, the probability that a replenishment order is outstanding is given by (10.1.9)

$$P(Y<r+1)=\frac{\mu}{\nu}\cdot P(X_1+X_2>0,Y=Q).$$

Summation of the global balance equations (10.1.1) over $n_1\in\mathbb{N}_0$ and $n_2\in\mathbb{N}_0$ leads to the following. Some of the changes are highlighted for reasons of clarity and comprehensibility.

$$\begin{aligned}
&\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\pi(n_1,n_2,k)\cdot\big((\lambda_1+p\lambda_2)\cdot 1_{\{0<k\le s\}}+(\lambda_1+\lambda_2)\cdot 1_{\{k>s\}}\\
&\qquad\qquad+\mu\cdot 1_{\{n_1+n_2>0\}}\cdot 1_{\{k>0\}}+\nu\cdot 1_{\{k\le r\}}\big)\\
&=\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\pi(n_1-1,n_2,k)\cdot\lambda_1\cdot 1_{\{n_1>0\}}\cdot 1_{\{k>0\}}\\
&\quad+\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\pi(n_1,n_2-1,k)\cdot p\,\lambda_2\cdot 1_{\{n_2>0\}}\cdot 1_{\{0<k\le s\}}
\end{aligned}$$

$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2 - 1, k) \cdot \lambda_2 \cdot 1_{\{n_2>0\}} \cdot 1_{\{k>s\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1 + 1, n_2, k + 1) \cdot \mu \cdot 1_{\{k<r+Q\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2 + 1, k + 1) \cdot \mu \cdot 1_{\{n_1=0\}} \cdot 1_{\{k<r+Q\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k - Q) \cdot \nu \cdot 1_{\{Q \le k \le r+Q\}}.$$

This is equivalent to

$$\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \big((\lambda_1 + p\lambda_2) \cdot 1_{\{0<k\le s\}} + (\lambda_1 + \lambda_2) \cdot 1_{\{k>s\}}$$
$$+\mu \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k\le r\}}\big) - \pi(0, 0, k) \cdot \mu \cdot 1_{\{k>0\}}$$
$$= \sum_{n_1=1}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1 - 1, n_2, k) \cdot \lambda_1 \cdot 1_{\{k>0\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=1}^{\infty} \pi(n_1, n_2 - 1, k) \cdot p\,\lambda_2 \cdot 1_{\{0<k\le s\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=1}^{\infty} \pi(n_1, n_2 - 1, k) \cdot \lambda_2 \cdot 1_{\{k>s\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1 + 1, n_2, k + 1) \cdot \mu \cdot 1_{\{k<r+Q\}}$$
$$+ \sum_{n_2=0}^{\infty} \pi(0, n_2 + 1, k + 1) \cdot \mu \cdot 1_{\{k<r+Q\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k - Q) \cdot \nu \cdot 1_{\{Q\le k\le r+Q\}}$$
$$= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \lambda_1 \cdot 1_{\{k>0\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot p\,\lambda_2 \cdot 1_{\{0<k\le s\}} + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \lambda_2 \cdot 1_{\{k>s\}}$$
$$+ \sum_{n_1=1}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k + 1) \cdot \mu \cdot 1_{\{k<r+Q\}}$$
$$+ \sum_{n_2=1}^{\infty} \pi(0, n_2, k + 1) \cdot \mu \cdot 1_{\{k<r+Q\}}$$
$$+ \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k - Q) \cdot \nu \cdot 1_{\{Q\le k\le r+Q\}}$$

$$
\begin{aligned}
&= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k)\cdot\lambda_1\cdot 1_{\{k>0\}}\\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k)\cdot p\,\lambda_2\cdot 1_{\{0<k\le s\}} + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k)\cdot\lambda_2\cdot 1_{\{k>s\}}\\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k+1)\cdot\mu\cdot 1_{\{k<r+Q\}} - \sum_{n_2=0}^{\infty}\pi(0,n_2,k+1)\cdot\mu\cdot 1_{\{k<r+Q\}}\\
&\quad + \sum_{n_2=0}^{\infty}\pi(0,n_2,k+1)\cdot\mu\cdot 1_{\{k<r+Q\}} - \pi(0,0,k+1)\cdot\mu\cdot 1_{\{k<r+Q\}}\\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k-Q)\cdot\nu\cdot 1_{\{Q\le k\le r+Q\}}.
\end{aligned}
$$

Cancelling on both sides the terms $\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\pi(n_1,n_2,k)\cdot(\lambda_1+p\,\lambda_2)\cdot 1_{\{0<k\le s\}}$ and $\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\pi(n_1,n_2,k)\cdot(\lambda_1+\lambda_2)\cdot 1_{\{k>s\}}$ yields

$$
\begin{aligned}
&\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k)\cdot\big(\mu\cdot 1_{\{k>0\}}+\nu\cdot 1_{\{k\le r\}}\big) - \pi(0,0,k)\cdot\mu\cdot 1_{\{k>0\}}\\
&= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k+1)\cdot\mu\cdot 1_{\{k<r+Q\}} - \pi(0,0,k+1)\cdot\mu\cdot 1_{\{k<r+Q\}}\\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k-Q)\cdot\nu\cdot 1_{\{Q\le k\le r+Q\}}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&P(X_1+X_2>0,Y=k)\cdot\mu\cdot 1_{\{k>0\}} + P(Y=k)\cdot\nu\cdot 1_{\{k\le r\}}\\
&= P(X_1+X_2>0,Y=k+1)\cdot\mu\cdot 1_{\{k<r+Q\}} + P(Y=k-Q)\cdot\nu\cdot 1_{\{Q\le k\le r+Q\}}.
\end{aligned}
$$

This leads to

$$
\nu\cdot P(Y=0) = \mu\cdot P(X_1+X_2>0,Y=1), \qquad k=0, \tag{10.1.12}
$$

$$
\begin{aligned}
\nu\cdot P(Y=k) &= \mu\cdot P(X_1+X_2>0,Y=k+1)\\
&\quad -\mu\cdot P(X_1+X_2>0,Y=k), \qquad 0<k\le r,
\end{aligned} \tag{10.1.13}
$$

$$
\begin{aligned}
&\mu\cdot P(X_1+X_2>0,Y=k)\\
&= \mu\cdot P(X_1+X_2>0,Y=k+1), \qquad r<k<Q,
\end{aligned} \tag{10.1.14}
$$

$$
\begin{aligned}
\nu\cdot P(Y=k) &= \mu\cdot P(X_1+X_2>0,Y=Q+k)\\
&\quad -\mu\cdot P(X_1+X_2>0,Y=Q+k+1), \qquad 0\le k<r,
\end{aligned} \tag{10.1.15}
$$

$$
\nu\cdot P(Y=r) = \mu\cdot P(X_1+X_2>0,Y=r+Q), \qquad k=r+Q. \tag{10.1.16}
$$

From (10.1.12) and (10.1.15) with $k = 0$ follows

$$\begin{aligned} & P(X_1 + X_2 > 0, Y = Q) \\ & = P(X_1 + X_2 > 0, Y = 1) + P(X_1 + X_2 > 0, Y = Q + 1). \end{aligned} \tag{10.1.17}$$

From (10.1.13) and (10.1.15) it can be deduced that for $k = 1, \dots, r-1$

$$\begin{aligned} & P(X_1 + X_2 > 0, Y = k) + P(X_1 + X_2 > 0, Y = Q + k) \\ & = P(X_1 + X_2 > 0, Y = k+1) + P(X_1 + X_2 > 0, Y = Q + k + 1) \end{aligned} \tag{10.1.18}$$

holds. Finally, for $k = r$ (10.1.13) and (10.1.16) yield

$$\begin{aligned} & P(X_1 + X_2 > 0, Y = r) + P(X_1 + X_2 > 0, Y = Q + r) \\ & = P(X_1 + X_2 > 0, Y = r + 1). \end{aligned} \tag{10.1.19}$$

Then (10.1.17)-(10.1.19) lead for $k = 1, \dots, r$ to

$$\begin{aligned} & P(X_1 + X_2 > 0, Y = Q) \\ & = P(X_1 + X_2 > 0, Y = k) + P(X_1 + X_2 > 0, Y = Q + k) \\ & = P(X_1 + X_2 > 0, Y = r + 1), \end{aligned} \tag{10.1.20}$$

which together with (10.1.9) yields (10.1.10).

Furthermore, it holds

$$\begin{aligned} P(X_1 + X_2 > 0, Y > 0) &= \sum_{k=1}^{r+Q} P(X_1 + X_2 > 0, Y = k) \\ &= \sum_{k=1}^{r} P(X_1 + X_2 > 0, Y = k) + \sum_{k=r+1}^{Q} P(X_1 + X_2 > 0, Y = k) \\ &\quad + \sum_{k=Q+1}^{r+Q} P(X_1 + X_2 > 0, Y = k) \\ &= \sum_{k=1}^{r} [P(X_1 + X_2 > 0, Y = k) + P(X_1 + X_2 > 0, Y = Q + k)] \\ &\quad + \sum_{k=r+1}^{Q} P(X_1 + X_2 > 0, Y = k) \\ &\overset{(10.1.14)}{=} \sum_{k=1}^{r} [P(X_1 + X_2 > 0, Y = k) + P(X_1 + X_2 > 0, Y = Q + k)] \\ &\quad + (Q - r) \cdot P(X_1 + X_2 > 0, Y = r + 1) \\ &\overset{(10.1.20)}{=} r \cdot P(X_1 + X_2 > 0, Y = r + 1) + (Q - r) \cdot P(X_1 + X_2 > 0, Y = r + 1) \\ &= Q \cdot P(X_1 + X_2 > 0, Y = r + 1) \end{aligned}$$

Therefore, the probability that a replenishment order is outstanding is given by (10.1.11)

$$P(Y < r+1) = \frac{\mu}{\nu \cdot Q} \cdot P(X_1 + X_2 > 0, Y > 0).$$

□

In the literature there are several different definitions of a unimodal function. The following definition is based on [Kei79, Definition 5.3B, p. 63].

Definition 10.1.7 (Weakly unimodal)**.** A distribution $P(X \leq x)$ with all support on the lattice of integers will be said to be weakly unimodal if there exists at least one integer $x^\star$ such that

$$P(X = x) \leq P(X = y) \quad \text{for all } x,\ y \text{ with } x \leq y \leq x^\star \text{ or } x^\star \leq y \leq x.$$

Proposition 10.1.8. *The joint probability density $P(X_1 + X_2 > 0, Y = k)$ is weakly unimodal*

$$\begin{aligned} P(X_1 + X_2 > 0, Y = k) &< P(X_1 + X_2 > 0, Y = k+1), && k = 1, \dots, r, \\ P(X_1 + X_2 > 0, Y = k) &= P(X_1 + X_2 > 0, Y = k+1), && k = r+1, \dots, Q-1, \\ P(X_1 + X_2 > 0, Y = k) &> P(X_1 + X_2 > 0, Y = k+1), && k = Q, \dots, Q+r-1. \end{aligned}$$

Hence, the conditional probability density $P(Y = k \mid X_1 + X_2 > 0)$ is weakly unimodal

$$\begin{aligned} P(Y = k \mid X_1 + X_2 > 0) &< P(Y = k+1 \mid X_1 + X_2 > 0), && k = 1, \dots, r, \\ P(Y = k \mid X_1 + X_2 > 0) &= P(Y = k+1 \mid X_1 + X_2 > 0), && k = r+1, \dots, Q-1, \\ P(Y = k \mid X_1 + X_2 > 0) &> P(Y = k+1 \mid X_1 + X_2 > 0), && k = Q, \dots, Q+r-1. \end{aligned}$$

Proof. For $k = 1, \dots, r$ holds from (10.1.13)

$$\frac{\nu}{\mu} \cdot P(Y = k) = P(X_1 + X_2 > 0, Y = k+1) - P(X_1 + X_2 > 0, Y = k).$$

Because of ergodicity it holds $P(Y = k) > 0$. Consequently, for $k = 1, \dots r$ yields

$$P(X_1 + X_2 > 0, Y = k) < P(X_1 + X_2 > 0, Y = k+1).$$

From (10.1.14) follows directly

$$P(X_1 + X_2 > 0, Y = k) = P(X_1 + X_2 > 0, Y = k+1), \qquad k = r+1, \dots, Q-1.$$

For $k = Q, \dots, Q+r-1$ holds from (10.1.15)

$$\frac{\nu}{\mu} \cdot P(Y = k-Q) = P(X_1 + X_2 > 0, Y = k) - P(X_1 + X_2 > 0, Y = k+1).$$

From ergodicity it follows that $P(Y = k-Q) > 0$. Consequently, for $k = Q, \dots, Q+r-1$ yields

$$P(X_1 + X_2 > 0, Y = k) > P(X_1 + X_2 > 0, Y = k+1).$$

The fact that $P(Y = k \mid X_1 + X_2 > 0) = \frac{P(X_1+X_2>0, Y=k)}{P(X_1+X_2>0)}$ and $P(X_1 + X_2 > 0) > 0$ hold implies the last three equations in the proposition. □

Remark 10.1.9. The statements of Proposition 10.1.6 and Proposition 10.1.8 exhibit an insensitivity property with respect to variation of the parameters of the system, more specifically it is independent of the threshold level s.

Remark 10.1.10. The results in this Section can be generalized in a direct way to the case of a system with C customer classes, where $\overline{C} = \{1, \ldots, C\}$ is the set of customer classes — the smaller the number, the higher the priority. Customers of type c have an arrival rate $\lambda_c > 0$, a priority parameter p_c, $0 \leq p_c \leq 1$, and a threshold level s_c, $c \in \overline{C}$.

10.2. Pure inventory system

In this section, we consider the case of zero service time, which is the version of our model in the classical inventory theory. The supply chain of interest is depicted in Figure 10.2.1 and consists of priority and ordinary customers, an inventory and a supplier.

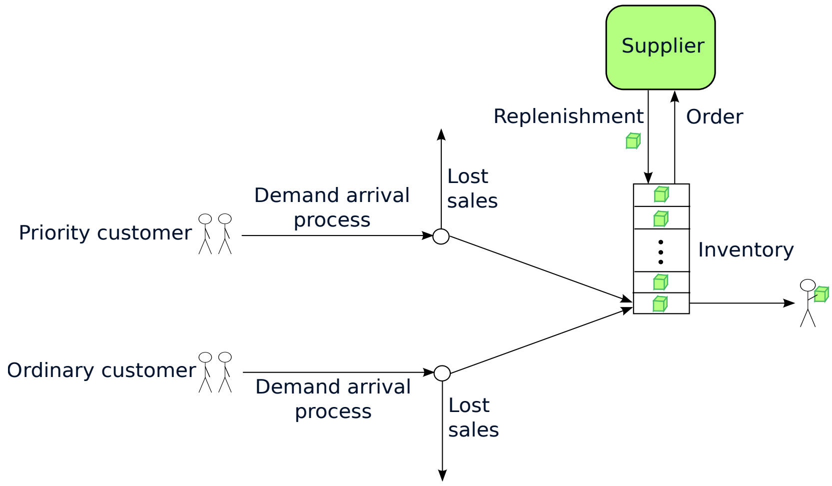


Figure 10.2.1.: The pure inventory system with two customer classes

There are two types of customers — priority customers and ordinary customers. $\overline{C} = \{1, 2\}$ is the set of these customer classes where 1 is the type of priority customers and 2 is the type of ordinary customers — the smaller the number, the higher the priority. According to two independent Poisson processes with different parameters the demands of each type of customer arrive one by one at the production system and require service. Priority customers arrive according to a Poisson process with rate $\lambda_1 > 0$ and ordinary customers arrive according to a Poisson process with rate $\lambda_2 > 0$.

Customers' arrivals are regulated by a flexible admission control with priority parameter p, $0 \leq p \leq 1$: If the inventory is depleted all arriving customers are rejected ("lost sales"). If the on-hand inventory is greater than a prescribed threshold level s, $0 < s < r + Q$, the customers of both classes are admitted to enter the system. If the on-hand inventory reaches or falls below the threshold level s, all priority customers still enter the system, but ordinary customers are allowed to enter only with probability p and are rejected with probability $1 - p$.

The inventory is depleted by an exogenous customer demand and each customer needs exactly one item from the inventory.

It is assumed that transmission times for orders are zero and that the transportation time between the production system and the inventory is negligible.
All inter-arrival times and replenishment lead times constitute an independent family of random variables.

An outside supplier replenishes raw material to the inventory according to the (r, Q)-policy. The replenishment lead time is exponentially distributed with parameter $\nu > 0$.
Let $Y(t)$ denote the on-hand inventory at time $t \geq 0$. Denote by $Y = (Y(t) : t \geq 0)$ the pure inventory process. Then, due to the usual independence and memoryless assumptions Y is a homogeneous strong Markov process. The state space of Y is

$$K = \{0, \ldots, r + Q\},$$

where $r + Q$ is the maximal size of the inventory, which depends on the replenishment policy.
The queueing-inventory process Y is irreducible. It can be shown analogously as in Appendix E on page 375 for the queueing-inventory system with base stock policy. From $|K| < \infty$ follows ergodicity (cf. [Ser13, Theorem 4.21]).

Definition 10.2.1. For the queueing-inventory process Z in a state space K, whose limiting distribution exists, we define

$$\theta := (\theta(k) : k \in K), \quad \theta(k) := \lim_{t \to \infty} P(Y(t) = k).$$

Proposition 10.2.2. *The inventory process $Y = (Y(t) : t \geq 0)$ has the following limiting and stationary distribution:*

(a) *If the threshold level is equal or greater than the reorder level, i.e. $r \leq s < r + Q$, then for $k = 1, 2, \ldots, r$ holds*

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{k-1} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0), \tag{10.2.1}$$

for $k = r + 1, \ldots, Q$ holds

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{r} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(k,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(k,s)=s\}})}\right) \cdot \theta(0), \tag{10.2.2}$$

for $k = Q + 1, \ldots, Q + r$ holds

$$\theta(k) = \left[\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{r} - \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{k-Q-1}\right] \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(k,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(k,s)=s\}})}\right) \cdot \theta(0), \tag{10.2.3}$$

and $\theta(0)$ can be calculated by the normalizing condition.

(b) *If the threshold level is not greater than the reorder level, i.e. $s < r$, then for $k = 1, 2, \ldots, s$ holds*

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{k-1} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0), \tag{10.2.4}$$

for $k = s+1, \ldots r$ holds

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{k-s-1} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0), \tag{10.2.5}$$

for $k = r+1, \ldots, Q$ holds

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{r-s} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0), \tag{10.2.6}$$

for $k = Q+1, \ldots, Q+s$ holds

$$\begin{aligned}\theta(k) = &\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{r-s} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0)\\ &- \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{k-Q-1} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0),\end{aligned} \tag{10.2.7}$$

for $k = Q+s+1, \ldots, Q+r$ holds

$$\begin{aligned}\theta(k) = &\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{r-s} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0)\\ &- \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{k-Q-s-1} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0),\end{aligned} \tag{10.2.8}$$

and $\theta(0)$ can be calculated by the normalizing condition.

Proof. **(a)** The global balance equations of the pure inventory system with $r \le s < r+Q$ are given as follows:

$$\theta(0) \cdot \nu = \theta(1) \cdot \big(\lambda_1 + p\,\lambda_2\big), \tag{10.2.9}$$

$$\theta(j) \cdot \big(\lambda_1 + p\,\lambda_2 + \nu\big) = \theta(j+1) \cdot \big(\lambda_1 + p\,\lambda_2\big), \qquad j = 1, \ldots, r-1, \tag{10.2.10}$$

$$\begin{aligned}&\theta(r) \cdot \big(\lambda_1 + p\,\lambda_2 + \nu\big)\\ &= \theta(r+1) \cdot \big(\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(r+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(r+1,s)=s\}})\big), \quad j = r,\end{aligned} \tag{10.2.11}$$

$$\begin{aligned}&\theta(j) \cdot \big(\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j,s)=s\}})\big)\\ &= \theta(j+1) \cdot \big(\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j+1,s)=s\}})\big),\\ &\qquad\qquad j = r+1, \ldots, Q-1,\end{aligned} \tag{10.2.12}$$

$$\theta(j) \cdot \big(\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j,s)=s\}})\big)$$
$$= \theta(j+1) \cdot \big(\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j+1,s)=s\}})\big) + \theta(j-Q) \cdot \nu, \quad j = Q, \ldots, Q+r-1, \quad (10.2.13)$$

$$\theta(Q+r) \cdot \big(\lambda_1 + \lambda_2\big) = \theta(r) \cdot \nu. \quad (10.2.14)$$

▶ (10.2.1) and (10.2.2) can be directly obtained by recursion from the above global balance equations.

• From (10.2.9) it follows
$$\theta(1) = \frac{\nu}{\lambda_1 + p\,\lambda_2} \cdot \theta(0)$$
and from (10.2.10) it follows for $k = 2, \ldots, r$
$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(k-1).$$
Hence, it holds (10.2.1) for $k = 1, 2, \ldots, r$
$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{k-1} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0).$$

• From (10.2.11) it follows
$$\begin{aligned}
\theta(r+1) &= \frac{\lambda_1 + p\,\lambda_2 + \nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(r+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(r+1,s)=s\}})} \cdot \theta(r) \\
&\overset{(10.2.1)}{=} \frac{\lambda_1 + p\,\lambda_2 + \nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(r+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(r+1,s)=s\}})} \cdot \left(1 + \frac{\nu}{\lambda_1 + p\lambda_2}\right)^{r-1} \\
&\quad \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0) \\
&= \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{r} \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(r+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(r+1,s)=s\}})}\right) \\
&\quad \cdot \theta(0)
\end{aligned}$$
and from (10.2.12) it follows for $k = r+2, \ldots, Q$
$$\theta(k) = \frac{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(k-1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(k-1,s)=s\}})}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(k,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(k,s)=s\}})} \cdot \theta(k-1).$$
Consequently, for $k = r+1, \ldots, Q$ holds (10.2.2)
$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{r} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(k,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(k,s)=s\}})}\right) \cdot \theta(0).$$

► Thus, it only has to be shown that the distribution (10.2.3) satisfies the global balance equations. Substitution of (10.2.1)-(10.2.3) into the global balance equations (10.2.13) and (10.2.14) directly leads to
for $j = Q$ holds

$$\begin{aligned}
&\theta(Q) \cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(Q,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(Q,s)=s\}}))\\
&= \theta(Q+1) \cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(Q+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(Q+1,s)=s\}})) + \theta(0) \cdot \nu\\
\overset{(10.2.2),\,(10.2.3)}{\Leftrightarrow}\quad &\left(1 + \frac{\nu}{\lambda_1 + p\lambda_2}\right)^r\\
&\cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(Q,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(Q,s)=s\}})}\right) \cdot \theta(0)\\
&\cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(Q,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(Q,s)=s\}}))\\
&= \left[\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^r - 1\right]\\
&\quad \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(Q+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(Q+1,s)=s\}})}\right) \cdot \theta(0)\\
&\quad \cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(Q+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(Q+1,s)=s\}})) + \theta(0) \cdot \nu\\
\Leftrightarrow\quad &\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^r = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^r - 1 + 1,
\end{aligned}$$

for $j = Q+1, \dots, r+Q-1$ holds

$$\begin{aligned}
&\theta(j) \cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j,s)=s\}}))\\
&= \theta(j+1) \cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j+1,s)=s\}})) + \theta(j-Q) \cdot \nu\\
\overset{(10.2.1),\,(10.2.3)}{\Leftrightarrow}\quad &\left[\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^r - \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{j-Q-1}\right]\\
&\cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j,s)=s\}})}\right) \cdot \theta(0)\\
&\cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j,s)=s\}}))\\
&= \left[\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^r - \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{j-Q}\right]\\
&\quad \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j+1,s)=s\}})}\right) \cdot \theta(0)\\
&\quad \cdot (\lambda_1 + p\,\lambda_2 \cdot 1_{\{\max(j+1,s)=s\}} + \lambda_2 \cdot (1 - 1_{\{\max(j+1,s)=s\}}))\\
&\quad + \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{j-Q-1} \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0) \cdot \nu
\end{aligned}$$

$$
\begin{aligned}
\Leftrightarrow \quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r}-\left(1+\frac{\nu}{\lambda_1+p\lambda_2}\right)^{j-Q-1} \\
&= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r}-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q} \\
&\quad + \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}\cdot\left(\frac{\nu}{\lambda_1+p\,\lambda_2}\right) \\
\Leftrightarrow \quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q} = \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right) \\
&= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q},
\end{aligned}
$$

for $j = Q + r$ holds

$$
\begin{aligned}
& \theta(Q+r)\cdot(\lambda_1+\lambda_2) = \theta(r)\cdot\nu \\
\overset{(10.2.1),\,(10.2.3)}{\Leftrightarrow} \quad & \left[\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r}-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r-1}\right]\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2) \\
&= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r-1}\left(\frac{\nu}{\lambda_1+p\,\lambda_2}\right)\cdot\theta(0)\cdot\nu \\
\Leftrightarrow \quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r}-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r-1} = \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r-1}\left(\frac{\nu}{\lambda_1+p\,\lambda_2}\right) \\
\Leftrightarrow \quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r} = \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r-1}\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right) \\
&= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{r}.
\end{aligned}
$$

(b) The global balance equations of the pure inventory system with $s < r < r + Q$ are given as follows

$$\theta(0)\cdot\nu = \theta(1)\cdot(\lambda_1+p\,\lambda_2), \tag{10.2.15}$$

$$\theta(j)\cdot(\lambda_1+p\,\lambda_2+\nu) = \theta(j+1)\cdot(\lambda_1+p\,\lambda_2), \qquad j = 1,\dots,s-1, \tag{10.2.16}$$

$$\theta(s)\cdot(\lambda_1+p\,\lambda_2+\nu) = \theta(s+1)\cdot(\lambda_1+\lambda_2), \qquad j = s, \tag{10.2.17}$$

$$\theta(j)\cdot(\lambda_1+\lambda_2+\nu) = \theta(j+1)\cdot(\lambda_1+\lambda_2), \qquad j = s+1,\dots,r, \tag{10.2.18}$$

$$\theta(j)\cdot(\lambda_1+\lambda_2) = \theta(j+1)\cdot(\lambda_1+\lambda_2), \qquad j = r+1,\dots,Q-1, \tag{10.2.19}$$

$$\theta(j) \cdot (\lambda_1 + \lambda_2) = \theta(j+1) \cdot (\lambda_1 + \lambda_2) + \theta(j-Q) \cdot \nu, \quad j = Q, \ldots, Q+r-1, \quad (10.2.20)$$

$$\theta(Q+r) \cdot (\lambda_1 + \lambda_2) = \theta(r) \cdot \nu. \quad (10.2.21)$$

► (10.2.4)-(10.2.6) can be directly obtained by recursion from the global balance equations.

• From (10.2.15) it follows

$$\theta(1) = \frac{\nu}{\lambda_1 + p\,\lambda_2} \cdot \theta(0)$$

and from (10.2.16) it follows for $k = 2, \ldots, s$

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(k-1).$$

Hence, it holds (10.2.4) for $k = 1, 2, \ldots, s$

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{k-1} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0).$$

• From (10.2.17) it follows

$$\begin{aligned}
\theta(s+1) &= \frac{\lambda_1 + p\,\lambda_2 + \nu}{\lambda_1 + \lambda_2} \cdot \theta(s) \\
&\overset{(10.2.4)}{=} \frac{\lambda_1 + p\,\lambda_2 + \nu}{\lambda_1 + \lambda_2} \cdot \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s-1} \cdot \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) \cdot \theta(0) \\
&= \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0)
\end{aligned}$$

and from (10.2.18) it follows for $k = s+2, \ldots, r+1$

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(k-1).$$

Consequently, for $k = s+1, \ldots, r+1$ holds (10.2.5)

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{k-s-1} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0).$$

• From (10.2.19) it follows for $k = r+1, \ldots, Q-1$

$$\theta(k) = \theta(k+1).$$

Hence, for $k = r+1, \ldots, Q$ it holds (10.2.6)

$$\theta(k) = \left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right)^{s} \cdot \left(1 + \frac{\nu}{\lambda_1 + \lambda_2}\right)^{r-s} \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2}\right) \cdot \theta(0).$$

► Thus, it only has to be shown that the distribution (10.2.7) and (10.2.8) satisfy the global balance equations. Substitution of (10.2.4)-(10.2.8) into the global balance equations (10.2.20) and (10.2.21) directly leads to
for $j = Q$ holds

$$\theta(Q)\cdot(\lambda_1+\lambda_2) = \theta(Q+1)\cdot(\lambda_1+\lambda_2) + \theta(0)\cdot\nu$$

$$\overset{(10.2.15),(10.2.20)}{\Leftrightarrow} \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)$$
$$= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)$$
$$-\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)+\theta(0)\cdot\nu$$

$$\Leftrightarrow \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}$$
$$= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s} - 1 + 1,$$

for $j = Q+1,\ldots,Q+s$ holds

$$\theta(j)\cdot(\lambda_1+\lambda_2) = \theta(j+1)\cdot(\lambda_1+\lambda_2)+\theta(j-Q)\cdot\nu$$

$$\overset{(10.2.16),(10.2.20)}{\Leftrightarrow} \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)$$
$$-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)$$
$$= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)$$
$$-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2)$$
$$+\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}\cdot\left(\frac{\nu}{\lambda_1+p\,\lambda_2}\right)\cdot\theta(0)\cdot\nu$$

$$\Leftrightarrow \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}$$
$$= \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s}\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q}$$
$$+\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}\cdot\left(\frac{\nu}{\lambda_1+p\,\lambda_2}\right)$$

$$\Leftrightarrow -\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}$$
$$= -\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q}+\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{j-Q-1}\cdot\left(\frac{\nu}{\lambda_1+p\,\lambda_2}\right)$$

$$\Leftrightarrow \qquad -1 = -\left(1 + \frac{\nu}{\lambda_1 + p\,\lambda_2}\right) + \left(\frac{\nu}{\lambda_1 + p\,\lambda_2}\right) = -1,$$

for $j = Q+s+1, \ldots, Q+r-1$ holds

$$\begin{aligned}
& \theta(j) \cdot (\lambda_1+\lambda_2) = \theta(j+1)\cdot(\lambda_1+\lambda_2) + \theta(j-Q)\cdot\nu \\
\overset{(10.2.16),\,(10.2.20)}{\Leftrightarrow} \quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \cdot \theta(0)\cdot(\lambda_1+\lambda_2) \\
& - \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s-1} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \cdot \theta(0)\cdot(\lambda_1+\lambda_2) \\
& = \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \cdot \theta(0)\cdot(\lambda_1+\lambda_2) \\
& \quad - \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \cdot \theta(0)\cdot(\lambda_1+\lambda_2) \\
& \quad + \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s-1} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \cdot \theta(0)\cdot\nu \\
\Leftrightarrow \quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s} \\
& - \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s-1} \\
& = \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s} \\
& \quad - \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s} \\
& \quad + \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s-1} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \\
\Leftrightarrow \quad & - \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s-1} \\
& = - \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s} \\
& \quad + \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^{s} \cdot \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{j-Q-s-1} \cdot \left(\frac{\nu}{\lambda_1+\lambda_2}\right) \\
\Leftrightarrow \quad & -1 = -\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right) + \left(\frac{\nu}{\lambda_1+\lambda_2}\right) = -1,
\end{aligned}$$

for $j = r + Q$ holds

$$\begin{aligned}
&\theta(Q+r)\cdot(\lambda_1+\lambda_2) = \theta(r)\cdot\nu \\
\overset{(10.2.18),\,(10.2.21)}{\Leftrightarrow}\quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^s\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2) \\
&-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^s\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s-1}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot(\lambda_1+\lambda_2) \\
&=\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^s\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s-1}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right)\cdot\theta(0)\cdot\nu \\
\Leftrightarrow\quad & \left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^s\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s} \\
&-\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^s\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s-1} \\
&=\left(1+\frac{\nu}{\lambda_1+p\,\lambda_2}\right)^s\cdot\left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)^{r-s-1}\cdot\left(\frac{\nu}{\lambda_1+\lambda_2}\right) \\
\Leftrightarrow\quad & \left(1+\frac{\nu}{\lambda_1+\lambda_2}\right)-1=\left(\frac{\nu}{\lambda_1+\lambda_2}\right) \\
\Leftrightarrow\quad & 1+\frac{\nu}{\lambda_1+\lambda_2}=1+\frac{\nu}{\lambda_1+\lambda_2}.
\end{aligned}$$

□

Remark 10.2.3. The equations (10.2.9)-(10.2.14) for $s = r$ have a similar structure as the global balance equations (1)-(6) of Liu and his coauthors in [LFW14, p. 181] of the single server inventory system with two classes of customers and flexible service policy. In their paper, the threshold level is equal to the reorder level.

Furthermore, the equations (10.2.9)-(10.2.14) for $p = 0$ and $r < s \leq Q$ have a similar structure as the global balance equations (A16)-(A22) of Isotupa in [Iso11, p. 16] of the single server inventory system with two classes of customers where the threshold level is above the reorder level.

The equations (10.2.15)-(10.2.21) for $p = 0$ and $0 < s \leq r$ have a similar structure as the global balance equations (A1)-(A7) of Isotupa in [Iso11, p. 14] of the single server inventory system with two classes of customers where the threshold level is not above the reorder level.

For $p = 1$, the model corresponds to the classic inventory system with (r, Q)-policy with arrival rate $\lambda_1 + \lambda_2$.

For $\lambda_2 \to 0$, the model corresponds to the classic inventory system with (r, Q)-policy with arrival rate λ_1.

10.3. Cost analysis

We consider the following cost structure for inventory, production and replenishment. The total costs consist of shortage costs $c_{ls,1}$ resp. $c_{ls,2}$ for each priority resp. ordinary customer that is lost, waiting costs $c_{w,1}$ resp. $c_{w,2}$ per unit of time for each priority resp. ordinary customer in the system (waiting or in service), capacity costs c_s per unit of time for providing inventory storage space (e.g. rent, insurance), holding costs c_h per unit of time for each unit that is kept on inventory. We assume that all of these costs per unit of time are positive.

Therefore, the cost function per unit of time in the respective states is

$$f_{r,Q,p}: \mathbb{N}_0 \times \mathbb{N}_0 \times \{0, 1, \dots, r+Q\} \longrightarrow \mathbb{R}_0^+$$

with

$$\begin{aligned} f_{r,Q,p}(n_1, n_2, k) = c_{w,1} \cdot n_1 + c_{w,2} \cdot n_2 + c_s \cdot (r+Q) + c_h \cdot k + c_{ls,1} \cdot \lambda_1 \cdot 1_{\{k=0\}} \\ + c_{ls,2} \cdot \left(\lambda_2 \cdot 1_{\{k=0\}} + (1-p) \cdot \lambda_2 \cdot 1_{\{0<k\leq s\}}\right). \end{aligned}$$

The asymptotic average costs for an ergodic system can be calculated as

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f_{r,Q,p}(Z(\omega,t))dt = \sum_{(n_1,n_2,k)} f_{r,Q,p}\left(n_1, n_2, k\right) \cdot \pi\left(n_1, n_2, k\right) =: \overline{f}(r, Q, p) \quad P\text{−}a.s.$$

Since we do not have a closed form expression for the stationary distribution of the queueing-inventory system, the cost optimization problem falls within the stochastic dynamic optimization (e.g. [HRS17], [Sen99]).

For the pure inventory system, the mixed integer optimization model as introduced in [LXC13, p. 1548] and [LFW14, pp. 182f.] can be used to provide an approximative solution. Moreover, to find the global optimum of the cost optimization problem for fixed p, a search algorithm as introduced by Liu and his coauthors [CZL12, Section 4.2, pp. 3085f.] can be used.

11. Production-inventory system with base stock policy

In this chapter, we study the production-inventory system with two types of customers, a flexible admission control and base stock policy, as described in Section 9.2.

Each item taken from the inventory results in a direct order sent to the supplier. This means, if a served customer departs from the system, an order for one item of the consumed raw material is placed at the supplier at this instant of time. The local base stock level $b \geq 2$ is the maximal size of the inventory. Note that there can be more than one outstanding order.

Customer arrivals are regulated by a flexible admission control with priority parameter p, $0 \leq p \leq 1$: If the inventory is depleted all arriving customers are rejected ("lost sales"). If the on-hand inventory is greater than a prescribed threshold level s, the customers of both classes are admitted to enter the system. If the on-hand inventory reaches or falls below the threshold level s, $0 < s < b$, all priority customers still enter the system but ordinary customers are allowed to enter only with probability p and are rejected with probability $1 - p$.

To obtain a Markovian process description of the integrated queueing-inventory system, we denote by $X_1(t)$ the number of priority customers present in the system at time $t \geq 0$, and by $X_2(t)$ the number of ordinary customers in the system at time $t \geq 0$ either waiting or in service (queue length). Since the customer in service will always be of the priority class when at least one priority customer is present, the value of the vector $(X_1(t), X_2(t))$ determines uniquely the type of the customer in service at time $t \geq 0$, if any. Moreover, by $Y(t)$ we denote the on-hand inventory at time $t \geq 0$.

We define the joint queueing-inventory process of this system by

$$Z = ((X_1(t), X_2(t), Y(t)) : t \geq 0).$$

Then, due to the usual independence and memoryless assumptions Z is a homogeneous Markov process. The state space of Z is

$$E = \left\{(n_1, n_2, k) : (n_1, n_2) \in \mathbb{N}_0^2,\ k \in \{0, \ldots, b\}\right\},$$

where b is the maximal size of the inventory.

In Section 11.1, we assume that the queueing-inventory process Z is ergodic to analyse first the properties of the stationary system. In Section 11.2, ergodicity is investigated in detail. In Section 11.3, we consider the case of zero service time.

11.1. Properties of the stationary system

In this section, we assume that the queueing-inventory process Z is ergodic.

The queueing-inventory process Z has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$ with the following transition rates for $(n_1, n_2, k) \in E$:

$$\begin{aligned} q((n_1,n_2,k);(n_1+1,n_2,k)) &= \lambda_1 \cdot 1_{\{k>0\}}, \\ q((n_1,n_2,k);(n_1,n_2+1,k)) &= p\,\lambda_2 \cdot 1_{\{0<k\le s\}} + \lambda_2 \cdot 1_{\{k>s\}}, \\ q((n_1,n_2,k);(n_1-1,n_2,k-1)) &= \mu \cdot 1_{\{n_1>0\}} \cdot 1_{\{k>0\}}, \\ q((n_1,n_2,k);(n_1,n_2-1,k-1)) &= \mu \cdot 1_{\{n_1=0\}} \cdot 1_{\{n_2>0\}} \cdot 1_{\{k>0\}}, \\ q((n_1,n_2,k);(n_1,n_2,k+1)) &= \nu \cdot 1_{\{k<b\}}. \end{aligned}$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\,(z;z) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\,(z;\tilde{z}) \qquad \forall z \in E.$$

Definition 11.1.1. For the queueing-inventory process Z on state space E, whose limiting distribution exists, we define

$$\pi := (\pi\,(n_1,n_2,k) : (n_1,n_2,k) \in E) \quad \text{by } \pi\,(n_1,n_2,k) := \lim_{t\to\infty} P\,(Z(t) = (n_1,n_2,k))\,.$$

The global balance equations $\pi \cdot \mathbf{Q} = \mathbf{0}$ of the ergodic queueing-inventory process Z are for $(n_1, n_2, k) \in E$ given by

$$\begin{aligned} &\pi(n_1,n_2,k) \cdot \big((\lambda_1 + p\,\lambda_2) \cdot 1_{\{0<k\le s\}} + (\lambda_1+\lambda_2) \cdot 1_{\{k>s\}} \\ &\qquad\qquad + \mu \cdot 1_{\{n_1+n_2>0\}} \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k<b\}}\big) \\ &= \pi(n_1-1,n_2,k) \cdot \lambda_1 \cdot 1_{\{n_1>0\}} \cdot 1_{\{k>0\}} \\ &\quad + \pi(n_1,n_2-1,k) \cdot p\,\lambda_2 \cdot 1_{\{n_2>0\}} \cdot 1_{\{0<k\le s\}} + \pi(n_1,n_2-1,k) \cdot \lambda_2 \cdot 1_{\{n_2>0\}} \cdot 1_{\{k>s\}} \\ &\quad + \pi(n_1+1,n_2,k+1) \cdot \mu \cdot 1_{\{k<b\}} + \pi(n_1,n_2+1,k+1) \cdot \mu \cdot 1_{\{n_1=0\}} \cdot 1_{\{k<b\}} \\ &\quad + \pi(n_1,n_2,k-1) \cdot \nu \cdot 1_{\{k>0\}}. \end{aligned} \tag{11.1.1}$$

Comparing the transition rates of this generator with the transition rates of the generator on page 212 of the production-inventory system with (r,Q)-policy, we see that they are identical if we choose $r = b-1$, $Q = 1$ and $r + Q = b$. Note, that this choice of parameters is not allowed in the (r,Q)-policy (cf. Definition on page 211) but this has no influence on the calculations in this section just involving the generator $\mathbf{Q}$. If we ignore the restriction $r < Q$ of the classic (r,Q)-policy, then most of the following results may be seen as special cases of the corresponding results in Section 10.1. We provide direct proofs here.

Let (X_1, X_2, Y) be a random variable that is distributed according to the queueing-inventory process in equilibrium. Then, Y is a random variable that is distributed according to the inventory process in equilibrium and X_1 resp. X_2 are random variables that are respectively distributed according to the queue length process of priority resp. ordinary customers in equilibrium.

Proposition 11.1.2. *For the queueing-inventory process holds the following equilibrium of probability flows*

$$\underbrace{P(Y>0)\cdot\lambda_1}_{\substack{\text{effective arrival rate}\\ \text{of priority customers}}} = \underbrace{P(X_1>0, Y>0)\cdot\mu}_{\substack{\text{effective departure rate}\\ \text{of priority customers}}}, \tag{11.1.2}$$

$$\underbrace{P(0<Y\le s)\cdot p\,\lambda_2 + P(Y>s)\cdot\lambda_2}_{\substack{\text{effective arrival rate}\\ \text{of ordinary customers}}} = \underbrace{P(X_1=0, X_2>0, Y>0)\cdot\mu}_{\substack{\text{effective departure rate}\\ \text{of ordinary customers}}}, \tag{11.1.3}$$

$$\begin{aligned}&\underbrace{P(Y>0)\cdot\lambda_1 + P(0<Y\le s)\cdot p\,\lambda_2 + P(Y>s)\cdot\lambda_2}_{\substack{\text{effective arrival rate}\\ \text{of customers}}},\\ &= \underbrace{P(X_1+X_2>0, Y>0)\cdot\mu}_{\substack{\text{effective departure rate}\\ \text{of customers}}}\end{aligned} \tag{11.1.4}$$

and

$$\begin{aligned}&P(X_1=n_1, Y>0)\cdot\lambda_1\\ &= P(X_1=n_1+1, Y>0)\cdot\mu, \qquad n_1\in\mathbb{N}_0,\end{aligned} \tag{11.1.5}$$

$$\begin{aligned}&P(X_2=n_2, 0<Y\le s)\cdot p\,\lambda_2 + P(X_2=n_2, Y>s)\cdot\lambda_2\\ &= P(X_1=0, X_2=n_2+1, Y>0)\cdot\mu, \qquad n_2\in\mathbb{N}_0,\end{aligned} \tag{11.1.6}$$

$$\begin{aligned}&P(X_1+X_2=n, 0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)\\ &+P(X_1+X_2=n, Y>s)\cdot(\lambda_1+\lambda_2)\\ &= P(X_1+X_2=n+1, 0<Y\le b)\cdot\mu, \qquad n\in\mathbb{N}_0.\end{aligned} \tag{11.1.7}$$

Remark 11.1.3. The effective departure rates are usually called throughput.

Proof. The equations can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a) on page 259.

For $n_1 \in \mathbb{N}_0$, equation (11.1.5) can be proven by a cut, which divides E into complementary sets according to the queue length of priority customers that is less than or equal to n_1 or greater than n_1, i.e. into the sets

$$\Big\{(m_1, m_2, k) : m_1 \in \{0, 1, \dots .n_1\},\ m_2 \in \mathbb{N}_0,\ k \in \{0, \dots, b\}\Big\},$$

$$\Big\{(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) : \widetilde{m}_1 \in \mathbb{N}_0 \setminus \{0, 1, \dots .n_1\},\ , \widetilde{m}_2 \in \mathbb{N}_0,\ \widetilde{k} \in \{0, \dots, b\}\Big\}, \quad n_1 \in \mathbb{N}_0.$$

Then, the following holds for $n_1 \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{m_1=0}^{n_1} \sum_{m_2=0}^{\infty} \sum_{k=0}^{b} \sum_{\widetilde{m}_1=n_1+1}^{\infty} \sum_{\widetilde{m}_2=0}^{\infty} \sum_{\widetilde{k}=0}^{b} \pi(m_1, m_2, k) \cdot q((m_1, m_2, k); (\widetilde{m}_1, \widetilde{m}_2, \widetilde{k})) \\
= &\sum_{\widetilde{m}_1=n_1+1}^{\infty} \sum_{\widetilde{m}_2=0}^{\infty} \sum_{\widetilde{k}=0}^{b} \sum_{m_1=0}^{n_1} \sum_{m_2=0}^{\infty} \sum_{k=0}^{b} \pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) \cdot q((\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}); (m_1, m_2, k)) \\
\Leftrightarrow \quad &\underbrace{\sum_{m_1=n_1}^{n_1} \sum_{m_2=0}^{\infty} \sum_{k=1}^{b} \pi(m_1, m_2, k) \cdot \lambda_1}_{=P(X_1=n_1, Y>0)\cdot\lambda_1} = \underbrace{\sum_{\widetilde{m}_1=n_1+1}^{n_1+1} \sum_{\widetilde{m}_2=0}^{\infty} \sum_{\widetilde{k}=1}^{b} \pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) \cdot \mu}_{=P(X_1=n_1+1, Y>0)\cdot\mu} .
\end{aligned}$$

Hence, for $n_1 \in \mathbb{N}_0$ holds (11.1.5)

$$P(X_1 = n_1, Y > 0) \cdot \lambda_1 = P(X_1 = n_1 + 1, Y > 0) \cdot \mu.$$

Summation of the equations (11.1.5) over $n_1 \in \mathbb{N}_0$ leads to

$$\sum_{n_1=0}^{\infty} P(X_1 = n_1, Y > 0) \cdot \lambda_1 = \sum_{n_1=0}^{\infty} P(X_1 = n_1 + 1, Y > 0) \cdot \mu,$$

which is equivalent to (11.1.2)

$$P(Y > 0) \cdot \lambda_1 = P(X_1 > 0, Y > 0) \cdot \mu.$$

For $n_2 \in \mathbb{N}_0$, equation (11.1.6) can be proven by a cut, which divides E into complementary sets according to the queue length of ordinary customers that is less than or equal to n_2 or greater than n_2, i.e. into the sets

$$\Big\{(m_1, m_2, k) : m_1 \in \mathbb{N}_0, m_2 \in \{0, 1, \ldots .n_2\},\ k \in \{0, \ldots, b\}\Big\},$$

$$\Big\{(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) : \widetilde{m}_1 \in \mathbb{N}_0, \widetilde{m}_2 \in \mathbb{N}_0 \setminus \{0, 1, \ldots .n_2\},\ \widetilde{k} \in \{0, \ldots, b\}\Big\}, \quad n_2 \in \mathbb{N}_0.$$

Then, the following holds for $n_2 \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{n_2}\sum_{k=0}^{b}\sum_{\widetilde{m}_1=0}^{\infty}\sum_{\widetilde{m}_2=n_2+1}^{\infty}\sum_{\widetilde{k}=0}^{b}\pi(m_1, m_2, k)\cdot q((m_1, m_2, k); (\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}))\\
&= \sum_{\widetilde{m}_1=0}^{\infty}\sum_{\widetilde{m}_2=n_2+1}^{\infty}\sum_{\widetilde{k}=0}^{b}\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{n_2}\sum_{k=0}^{b}\pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k})\cdot q((\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}); (m_1, m_2, k))\\
\Leftrightarrow\quad &\underbrace{\sum_{m_1=0}^{\infty}\sum_{m_2=n_2}^{n_2}\sum_{k=1}^{s}\pi(m_1, m_2, k)\cdot p\,\lambda_2}_{=P(X_2=n_2, 0<Y\le s)\cdot p\,\lambda_2} + \underbrace{\sum_{m_1=0}^{\infty}\sum_{m_2=n_2}^{n_2}\sum_{k=s+1}^{b}\pi(m_1, m_2, k)\cdot\lambda_2}_{=P(X_2=n_2, Y>s)\cdot\lambda_2}\\
&= \underbrace{\sum_{\widetilde{m}_1=0}^{0}\sum_{\widetilde{m}_2=n_2+1}^{n_2+1}\sum_{\widetilde{k}=1}^{b}\pi(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k})\cdot\mu}_{=P(X_1=0, X_2=n_2+1, Y>0)\cdot\mu}.
\end{aligned}$$

Thus, for $n_2 \in \mathbb{N}_0$ holds (11.1.6)

$$\begin{aligned}
&P(X_2 = n_2, 0 < Y \le s)\cdot p\,\lambda_2 + P(X_2 = n_2, Y > s)\cdot\lambda_2\\
&= P(X_1 = 0, X_2 = n_2 + 1, Y > 0)\cdot\mu.
\end{aligned}$$

Summation of the equations (11.1.6) over $n_2 \in \mathbb{N}_0$ leads to

$$\begin{aligned}
&\sum_{n_2=0}^{\infty} P(X_2 = n_2, 0 < Y \le s)\cdot p\,\lambda_2 + \sum_{n_2=0}^{\infty} P(X_2 = n_2, Y > s)\cdot\lambda_2\\
&= \sum_{n_2=0}^{\infty} P(X_1 = 0, X_2 = n_2 + 1, Y > 0)\cdot\mu,
\end{aligned}$$

which is equivalent to (11.1.3)

$$P(0 < Y \le s)\cdot p\,\lambda_2 + P(Y > s)\cdot\lambda_2 = P(X_1 = 0, X_2 > 0, Y > 0)\cdot\mu.$$

For $n \in \mathbb{N}_0$, equation (11.1.7) can be proven by a cut, which divides E into complementary sets according to the size of the total queue length that is less than or equal to n or greater than n, i.e. into the sets

$$\Big\{(m_1, m_2, k) : m_1 \in \mathbb{N}_0,\ m_2 \in \mathbb{N}_0,\ (m_1+m_2) \in \{0,1,\ldots,n\},\ k \in \{0,\ldots,b\}\Big\},$$

$$\Big\{(\widetilde{m}_1, \widetilde{m}_2, \widetilde{k}) : \widetilde{m}_1 \in \mathbb{N}_0,\ \widetilde{m}_2 \in \mathbb{N}_0,\ (\widetilde{m}_1+\widetilde{m}_2) \in \mathbb{N}_0 \setminus \{0,1,\ldots,n\},\ \widetilde{k} \in \{0,\ldots,b\}\Big\},\ n \in \mathbb{N}_0.$$

Then, the following holds for $n \in \mathbb{N}_0$

$$\begin{aligned}
&\sum_{m_1+m_2=0}^{n} \sum_{k=0}^{b} \sum_{\widetilde{m}_1+\widetilde{m}_2=n+1}^{\infty} \sum_{\widetilde{k}=0}^{b} \pi(m_1,m_2,k) \cdot q((m_1,m_2,k);(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k}))\\
= &\sum_{\widetilde{m}_1+\widetilde{m}_2=n+1}^{\infty} \sum_{\widetilde{k}=0}^{b} \sum_{m_1+m_2=0}^{n} \sum_{k=0}^{b} \pi(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k}) \cdot q((\widetilde{m}_1,\widetilde{m}_2,\widetilde{k});(m_1,m_2,k))\\
\Leftrightarrow \quad &\underbrace{\sum_{m_1+m_2=n} \sum_{k=1}^{s} \pi(m_1,m_2,k) \cdot (\lambda_1 + p\,\lambda_2)}_{=P(X_1+X_2=n, 0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)} + \underbrace{\sum_{m_1+m_2=n} \sum_{k=s+1}^{b} \pi(m_1,m_2,k) \cdot (\lambda_1+\lambda_2)}_{=P(X_1+X_2=n, Y>s)\cdot(\lambda_1+\lambda_2)}\\
= &\underbrace{\sum_{\widetilde{m}_1+\widetilde{m}_2=n+1}^{n+1} \sum_{\widetilde{k}=1}^{b} \pi(\widetilde{m}_1,\widetilde{m}_2,\widetilde{k}) \cdot \mu}_{=P(X_1+X_2=n+1, Y>0)\cdot\mu}\,.
\end{aligned}$$

Thus, for $n \in \mathbb{N}_0$ holds (11.1.7)

$$\begin{aligned}
&P(X_1+X_2=n, 0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)\\
&+P(X_1+X_2=n, Y>s)\cdot(\lambda_1+\lambda_2)\\
=\ &P(X_1+X_2=n+1, Y>0)\cdot\mu.
\end{aligned}$$

Summation of the equations (11.1.7) over $n \in \mathbb{N}_0$ leads to

$$\begin{aligned}
&\sum_{n=0}^{\infty} P(X_1+X_2=n, 0<Y\le s)\cdot(\lambda_1+p\,\lambda_2)\\
&+\sum_{n=0}^{\infty} P(X_1+X_2=n, Y>s)\cdot(\lambda_1+\lambda_2)\\
=&\sum_{n=0}^{\infty} P(X_1+X_2=n+1, Y>0)\cdot\mu,
\end{aligned}$$

which is equivalent to (11.1.4)

$$\begin{aligned}
&P(0<Y\le s)\cdot(\lambda_1+p\,\lambda_2) + P(Y>s)\cdot(\lambda_1+\lambda_2)\\
&= P(X_1+X_2>0, Y>0)\cdot\mu\\
\Leftrightarrow \quad &P(Y>0)\cdot\lambda_1 + P(0<Y\le s)\cdot p\,\lambda_2 + P(Y>s)\cdot\lambda_2\\
&= P(X_1+X_2>0, Y>0)\cdot\mu.
\end{aligned}$$

□

Remark 11.1.4.

(a) From (11.1.5) follows for $n_1 \in \mathbb{N}_0$

$$P(X_1 = n_1, Y > 0) = P(X_1 = 0, Y > 0) \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1}. \tag{11.1.8}$$

(b) In an $M/M/c$ queue with two priority classes under a preemptive priority discipline without inventory, the priority customers form a classic $M/M/c$ queue (cf. [WBSW15]). However, our queueing-inventory system with $c = 1$ (one server) is more difficult to solve since two customer classes have to share the same inventory and therefore, the priority customers do not form a classic $M/M/1$ queue.

(c) Rearranging (11.1.4) shows that the probability that the inventory is not depleted is given by

$$\begin{aligned} P(Y > 0) = {} & P(X_1 + X_2 > 0, Y > 0) \cdot \frac{\mu}{\lambda_1} \\ & - P(0 < Y \leq s) \cdot \frac{p\,\lambda_2}{\lambda_1} - P(Y > s) \cdot \frac{\lambda_2}{\lambda_1} \end{aligned}$$

and from (11.1.2) follows

$$P(Y > 0) = P(X_1 > 0, Y > 0) \cdot \frac{\mu}{\lambda_1}.$$

Corollary 11.1.5. *For the conditional distribution of the queue length process of priority customers conditioned on $\{Y > 0\}$ holds for $n_1 \in \mathbb{N}_0$*

$$P\left(X_1 = n_1 | Y > 0\right) = P\left(X_1 = 0 | Y > 0\right) \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1}$$

with

$$P\left(X_1 = 0 | Y > 0\right) = \left(1 - \frac{\lambda_1}{\mu}\right).$$

This shows that the conditional queue length process of priority customers under the condition that the inventory is not empty has in equilibrium the same structure as a birth-and-death process with birth-rates λ_1 and death-rates μ.

Proof. Equation (11.1.8) implies for $n \in \mathbb{N}_0$

$$\begin{aligned} P\left(X_1 = n_1 | Y > 0\right) &= \frac{P(X_1 = n_1, Y > 0)}{P\left(Y > 0\right)} \\ &\overset{(11.1.8)}{=} \frac{P(X_1 = 0, Y > 0)}{P\left(Y > 0\right)} \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1} \\ &= P\left(X_1 = 0 | Y > 0\right) \cdot \left(\frac{\lambda_1}{\mu}\right)^{n_1} \end{aligned}$$

and the normalizing condition leads to

$$P\left(X_1 = 0 | Y > 0\right) = \left(\sum_{n=1}^{\infty} \left(\frac{\lambda_1}{\mu}\right)^{n_1}\right)^{-1} = \left(1 - \frac{\lambda_1}{\mu}\right).$$

□

Proposition 11.1.6. *The probability that a replenishment order is outstanding fulfils the following equalities:*

(a)
$$P(Y < b) = \frac{\mu}{\nu} \cdot P(X_1 + X_2 > 0, Y > 0). \tag{11.1.9}$$

(b)
$$P(Y = b-1) = \frac{\mu}{\nu} \cdot P(X_1 + X_2 > 0, Y = b).$$

Proof. **(a)** Summation of the global balance equations (11.1.1) over $n_1 \in \mathbb{N}_0$ and $n_2 \in \mathbb{N}_0$ leads to the following. Some of the changes are highlighted for reasons of clarity and comprehensibility.

$$\begin{aligned}
&\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k)\cdot\Big((\lambda_1 + p\,\lambda_2)\cdot 1_{\{0<k\le s\}} + (\lambda_1+\lambda_2)\cdot 1_{\{k>s\}} \\
&\qquad\qquad\qquad\qquad + \mu\cdot 1_{\{n_1+n_2>0\}}\cdot 1_{\{k>0\}} + \nu\cdot 1_{\{k<b\}}\Big) \\
&= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1-1,n_2,k)\cdot\lambda_1\cdot 1_{\{n_1>0\}}\cdot 1_{\{k>0\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2-1,k)\cdot p\,\lambda_2\cdot 1_{\{n_2>0\}}\cdot 1_{\{0<k\le s\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2-1,k)\cdot \lambda_2\cdot 1_{\{n_2>0\}}\cdot 1_{\{k>s\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1+1,n_2,k+1)\cdot\mu\cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2+1,k+1)\cdot\mu\cdot 1_{\{n_1=0\}}\cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k-1)\cdot\nu\cdot 1_{\{k>0\}}.
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
&\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k) \\
&\cdot\big((\lambda_1 + p\,\lambda_2)\cdot 1_{\{0<k\le s\}} + (\lambda_1+\lambda_2)\cdot 1_{\{k>s\}} + \mu\cdot 1_{\{k>0\}} + \nu\cdot 1_{\{k<b\}}\big) \\
&-\pi(0,0,k)\cdot\mu\cdot 1_{\{k>0\}} \\
&= \sum_{n_1=1}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1-1,n_2,k)\cdot\lambda_1\cdot 1_{\{k>0\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=1}^{\infty} \pi(n_1,n_2-1,k)\cdot p\,\lambda_2\cdot 1_{\{0<k\le s\}} + \sum_{n_1=0}^{\infty}\sum_{n_2=1}^{\infty} \pi(n_1,n_2-1,k)\cdot\lambda_2\cdot 1_{\{k>s\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1+1,n_2,k+1)\cdot\mu\cdot 1_{\{k<b\}} + \sum_{n_2=0}^{\infty} \pi(0,n_2+1,k+1)\cdot\mu\cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \pi(n_1,n_2,k-1)\cdot\nu\cdot 1_{\{k>0\}}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \lambda_1 \cdot 1_{\{k>0\}} \\
&\quad + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot p\,\lambda_2 \cdot 1_{\{0<k\leq s\}} + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \lambda_2 \cdot 1_{\{k>s\}} \\
&\quad + \sum_{n_1=1}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k+1) \cdot \mu \cdot 1_{\{k<b\}} + \sum_{n_2=1}^{\infty} \pi(0, n_2, k+1) \cdot \mu \cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k-1) \cdot \nu \cdot 1_{\{k>0\}} \\
&= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \lambda_1 \cdot 1_{\{k>0\}} \\
&\quad + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot p\,\lambda_2 \cdot 1_{\{0<k\leq s\}} + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \lambda_2 \cdot 1_{\{k>s\}} \\
&\quad + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k+1) \cdot \mu \cdot 1_{\{k<b\}} - \sum_{n_2=0}^{\infty} \pi(0, n_2, k+1) \cdot \mu \cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_2=0}^{\infty} \pi(0, n_2, k+1) \cdot \mu \cdot 1_{\{k<b\}} - \pi(0,0,k+1) \cdot \mu \cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k-1) \cdot \nu \cdot 1_{\{k>0\}}.
\end{aligned}
$$

Cancelling on both sides the terms $\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot (\lambda_1 + p\,\lambda_2) \cdot 1_{\{0<k\leq s\}}$ and $\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot (\lambda_1 + \lambda_2) \cdot 1_{\{k>s\}}$ yields

$$
\begin{aligned}
&\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k) \cdot \big(\mu \cdot 1_{\{k>0\}} + \nu \cdot 1_{\{k<b\}}\big) \\
&-\pi(0,0,k) \cdot \mu \cdot 1_{\{k>0\}} \\
&= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k+1) \cdot \mu \cdot 1_{\{k<b\}} - \pi(0,0,k+1) \cdot \mu \cdot 1_{\{k<b\}} \\
&\quad + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \pi(n_1, n_2, k-1) \cdot \nu \cdot 1_{\{k>0\}}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&P(X_1 + X_2 > 0, Y = k) \cdot \mu \cdot 1_{\{k>0\}} + P(Y = k) \cdot \nu \cdot 1_{\{k<b\}} \\
&= P(X_1 + X_2 > 0, Y = k+1) \cdot \mu \cdot 1_{\{k<b\}} + P(Y = k-1) \cdot \nu \cdot 1_{\{k>0\}},
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
&P(Y = k) \cdot \nu \cdot 1_{\{k<b\}} - P(Y = k-1) \cdot \nu \cdot 1_{\{k>0\}} \\
&= P(X_1 + X_2 > 0, Y = k+1) \cdot \mu \cdot 1_{\{k<b\}} - P(X_1 + X_2 > 0, Y = k) \cdot \mu \cdot 1_{\{k>0\}}.
\end{aligned}
$$

This leads to

$$\nu \cdot P(Y=0) = \mu \cdot P(X_1 + X_2 > 0, Y = 1), \tag{11.1.10}$$

$$\begin{aligned}&\nu \cdot P(Y=k) - \nu \cdot P(Y=k-1) \\ &= \mu \cdot P(X_1 + X_2 > 0, Y = k+1) - \mu \cdot P(X_1 + X_2 > 0, Y = k),\ 1 \le k \le b-1,\end{aligned} \tag{11.1.11}$$

$$\nu \cdot P(Y=b-1) = \mu \cdot P(X_1 + X_2 > 0, Y = b). \tag{11.1.12}$$

(11.1.13)

From (11.1.10) and (11.1.11) with $k=1$ follows

$$\nu \cdot P(Y=1) = \mu \cdot P(X_1 + X_2 > 0, Y = 2) \tag{11.1.14}$$

because

$$\begin{aligned}&\nu \cdot P(Y=1) - \nu \cdot P(Y=0) \\ &\overset{(11.1.11)}{=} \mu \cdot P(X_1 + X_2 > 0, Y = 2) - \mu \cdot P(X_1 + X_2 > 0, Y = 1) \\ \overset{(11.1.10)}{\Leftrightarrow}\quad &\nu \cdot P(Y=1) - \mu \cdot P(X_1 + X_2 > 0, Y = 1) \\ &= \mu \cdot P(X_1 + X_2 > 0, Y = 2) - \mu \cdot P(X_1 + X_2 > 0, Y = 1) \\ \Leftrightarrow\quad &\nu \cdot P(Y=1) = \mu \cdot P(X_1 + X_2 > 0, Y = 2).\end{aligned}$$

Furthermore, with $k=2$ follows from (11.1.11)

$$\begin{aligned}&\nu \cdot P(Y=2) - \nu \cdot P(Y=1) \\ &= \mu \cdot P(X_1 + X_2 > 0, Y = 3) - \mu \cdot P(X_1 + X_2 > 0, Y = 2) \\ \overset{(11.1.14)}{\Leftrightarrow}\quad &\nu \cdot P(Y=2) - \mu \cdot P(X_1 + X_2 > 0, Y = 2) \\ &= \mu \cdot P(X_1 + X_2 > 0, Y = 3) - \mu \cdot P(X_1 + X_2 > 0, Y = 2) \\ \Leftrightarrow\quad &\nu \cdot P(Y=2) = \mu \cdot P(X_1 + X_2 > 0, Y = 3)\end{aligned}$$

and so on. Similarly we obtain for $k = 1, \dots, b-1$

$$\nu \cdot P(Y=k) = \mu \cdot P(X_1 + X_2 > 0, Y = k+1). \tag{11.1.15}$$

Hence,

$$\begin{aligned}P(X_1 + X_2 > 0, Y > 0) &= \sum_{k=1}^{b} P(X_1 + X_2 > 0, Y = k) \\ &= P(X_1 + X_2 > 0, Y = 1) + \sum_{k=2}^{b} P(X_1 + X_2 > 0, Y = k) \\ &\overset{\substack{(11.1.10),\\(11.1.15)}}{=} \sum_{k=1}^{b} \frac{\nu}{\mu} \cdot P(Y=k-1) = \sum_{k=0}^{b-1} \frac{\nu}{\mu} \cdot P(Y=k) = \frac{\nu}{\mu} \cdot P(Y<b).\end{aligned}$$

Therefore, the probability that a replenishment order is outstanding is given by

$$P(Y < b) = \frac{\mu}{\nu} \cdot P(X_1 + X_2 > 0, Y > 0).$$

(b) The equation can be proven by the cut-criterion for positive recurrent processes, which is presented in Theorem A.1.1(a) on page 259. For $b \geq 2$, it can be proven by a cut, which divides E into complementary sets according to the size of the inventory that is less than or equal to $b-1$ or greater than $b-1$, i.e. into the sets

$$\Big\{(n_1, n_2, k) : (n_1, n_2) \in \mathbb{N}_0^2,\ k \in \{0, \ldots, b-1\}\Big\},$$

$$\Big\{(\widetilde{n}_1, \widetilde{n}_2, \widetilde{k}) : (\widetilde{n}_1, \widetilde{n}_2) \in \mathbb{N}_0^2,\ \widetilde{k} \in \{b\}\Big\}, \quad b \geq 2.$$

Then, the following holds for $b \geq 2$

$$\begin{aligned}
&\sum_{n_1+n_2=0}^{\infty} \sum_{k=0}^{b-1} \sum_{\widetilde{n}_1+\widetilde{n}_2=0}^{\infty} \sum_{\widetilde{k}=b}^{b} \pi(n_1, n_2, k) \cdot q((n_1, n_2, k); (\widetilde{n}_1, \widetilde{n}_2, \widetilde{k})) \\
= &\sum_{\widetilde{n}_1+\widetilde{n}_2=0}^{\infty} \sum_{\widetilde{k}=b}^{b} \sum_{n_1+n_2=0}^{\infty} \sum_{k=0}^{b-1} \pi(\widetilde{n}_1, \widetilde{n}_2, \widetilde{k}) \cdot q((\widetilde{n}_1, \widetilde{n}_2, \widetilde{k}); (n_1, n_2, k)) \\
\Leftrightarrow \quad &\underbrace{\sum_{n_1+n_2=0}^{\infty} \pi(n_1, n_2, b-1) \cdot \nu}_{=P(Y=b-1)\cdot\nu} = \underbrace{\sum_{\widetilde{n}_1+\widetilde{n}_2=1}^{\infty} \pi(\widetilde{n}_1, \widetilde{n}_2, b) \cdot \mu}_{=P(X_1+X_2>0, Y=b)\cdot\mu}.
\end{aligned}$$

Thus, for $b \geq 2$ follows

$$P(Y = b-1) = \frac{\mu}{\nu} \cdot P(X_1 + X_2 > 0, Y = b).$$

□

Proposition 11.1.7. *For the joint probability density $P(X_1 + X_2 > 0, Y < k)$ holds*

$$P(X_1 + X_2 > 0, Y < k) < P(X_1 + X_2 > 0, Y < k+1), \qquad k = 0, \ldots, b-1.$$

Hence, for the conditioned probability density $P(Y < k \mid X_1 + X_2 > 0)$ holds

$$P(Y < k \mid X_1 + X_2 > 0) < P(Y < k+1 \mid X_1 + X_2 > 0), \qquad k = 1, \ldots, b-1.$$

Proof. For $k = 0, \ldots, b-1$ holds from (11.1.10) and (11.1.15)

$$\frac{\nu}{\mu} \cdot P(Y = k) = P(X_1 + X_2 > 0, Y = k+1).$$

Hence, from $P(Y < k) < P(Y < k+1)$ follows directly for $k = 0, \ldots b-1$

$$P(X_1 + X_2 > 0, Y < k) < P(X_1 + X_2 > 0, Y < k+1).$$

The fact that $P(Y < k \mid X_1 + X_2 > 0) = \frac{P(X_1+X_2>0, Y<k)}{P(X_1+X_2>0)}$ and $P(X_1 + X_2 > 0) > 0$ hold implies

$$P(Y < k \mid X_1 + X_2 > 0) < P(Y < k+1 \mid X_1 + X_2 > 0), \qquad k = 1, \ldots, b-1.$$

□

Remark 11.1.8. The statements of Proposition 11.1.6 and Proposition 11.1.7 exhibit an insensitivity property with respect to variation of the parameters of the system, more specifically it is independent of the threshold level s.

Remark 11.1.9. The results in this Section can be generalized in a direct way to the case of a system with C customer classes, where $\overline{C} = \{1, \dots, C\}$ is the set of customer classes — the smaller the number, the higher the priority. Customers of type c have an arrival rate $\lambda_c > 0$, a priority parameter p_c ($0 \leq p_c \leq 1$) and a threshold level s_c, $c \in \overline{C}$.

Remark 11.1.10. In the models of Isotupa [Iso15] the replenishment rate depends on the number of pending orders.

We can extend our model so that the replenishment lead time depends on the number of orders at the supplier. If there are $b-k > 0$ orders present at the supplier, the intensity of the replenishment lead time is $\nu(b-k) > 0$. Proposition 11.1.2, Remark 11.1.4, Remark 11.1.4 and Proposition 11.1.7 apply to this extension. In Proposition 11.1.6, we obtain for the extension

$$\sum_{k=0}^{b-1} P(Y = k) \cdot \nu(b-k) = \mu \cdot P(X_1 + X_2 > 0, Y > 0),$$

$$P(Y = b-1) = \frac{\mu}{\nu(1)} \cdot P(X_1 + X_2 > 0, Y = b).$$

11.2. Ergodicity

If our queueing-inventory process Z is irreducible and positive recurrent, then it is ergodic. The proof for irreducibility is exemplarily for $p = 1$ presented in Appendix E on page 375.

We will utilize the Foster-Lyapunov stability criterion, which is presented in Theorem A.1.2 on page 260. Another approach to show positive recurrence is by matrix analytic methods for level-dependent and level-independent quasi-birth-and-death processes [LR99]. A particular feature of the model is that the state space is infinite in two dimensions — the dimension of the priority customers and the dimension of the ordinary customers. Therefore, it is not obvious which dimension should play the role of the levels.

We obtain the following result by the Foster-Lyapunov stability criterion.

Theorem 11.2.1. *The queueing-inventory process Z is ergodic if $\lambda_1 + \lambda_2 < \mu$ and $\nu < \mu$.*

Before we prove this theorem above, we prove the following lemma.

Lemma 11.2.2. *The following conditions for some $\varepsilon > 0$ and some $C > 0$*

$$C \cdot (\nu - \mu) \leq -\varepsilon, \tag{11.2.1}$$

$$\lambda_1 + p\,\lambda_2 - \mu + C \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1 \right) + \nu \cdot \left(\frac{\nu}{\mu} - 1 \right) \right) \leq -\varepsilon, \tag{11.2.2}$$

$$\lambda_1 + \lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu} \right)^s \left(\mu \cdot \left(\frac{\mu}{\nu} - 1 \right) + \nu \cdot \left(\frac{\nu}{\mu} - 1 \right) \right) \leq -\varepsilon \tag{11.2.3}$$

imply

$$\lambda_1 + p\,\lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right) \leq -\varepsilon, \quad k = 1, \ldots, s, \tag{11.2.4}$$

$$\lambda_1 + \lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right) \leq -\varepsilon, \quad k = s+1, \ldots, b-1. \tag{11.2.5}$$

Proof. From condition (11.2.1) and $C > 0$

$$C \cdot (\nu - \mu) \leq -\varepsilon,$$

it follows

$$\mu > \nu. \tag{11.2.6}$$

▶ From condition (11.2.2) it follows condition (11.2.4)

$$\lambda_1 + p\,\lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right) \leq -\varepsilon, \qquad k = 1, \ldots, s,$$

since the worst case is $k = 1$. More precisely, from condition (11.2.6) follows

$$1 > \left(\frac{\nu}{\mu}\right) > \left(\frac{\nu}{\mu}\right)^2 > \left(\frac{\nu}{\mu}\right)^3 > \ldots$$

and

$$\begin{aligned} & \frac{\mu}{\nu} > \frac{\nu}{\mu} \\ \Leftrightarrow \quad & \mu \cdot \left(\frac{\mu - \nu}{\nu}\right) > \nu \cdot \left(\frac{\mu - \nu}{\mu}\right) \\ \Rightarrow \quad & \mu \cdot \left(\frac{\mu - \nu}{\nu}\right) + \nu \cdot \left(\frac{\nu - \mu}{\mu}\right) > 0 \\ \Leftrightarrow \quad & \mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right) > 0. \end{aligned}$$

▶ From condition (11.2.3) it follows condition (11.2.5)

$$\lambda_1 + \lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right) \leq -\varepsilon, \qquad k = s+1, \ldots, b-1,$$

since the worst case is $k = s + 1$. More precisely, from condition (11.2.6) follows

$$\left(\frac{\nu}{\mu}\right)^s > \left(\frac{\nu}{\mu}\right)^{s+1} > \left(\frac{\nu}{\mu}\right)^{s+2} > \ldots$$

and as shown above

$$\frac{\mu}{\nu} > \frac{\nu}{\mu} \qquad \Leftrightarrow \qquad \mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right) > 0.$$

□

Proof of Theorem 11.2.1. The positive recurrence can be shown by the Foster-Lyapunov stability criterion. We will show that $\mathcal{L} : E \to \mathbb{R}_0^+$ is a Lyapunov function with

$$\mathcal{L}(n_1, n_2, k) = n_1 + n_2 + \alpha(k) \tag{11.2.7}$$

where

$$\alpha(k) = C \cdot \left(\frac{\nu}{\mu}\right)^k \cdot \frac{\mu}{\nu}, \qquad k = 0, 1, \ldots, b, \tag{11.2.8}$$

with

$$C = \frac{\mu - (\lambda_1 + \lambda_2)}{2 \cdot \widetilde{C}}, \tag{11.2.9}$$

where

$$\begin{aligned}\widetilde{C} = \max\Bigg\{&(\mu - \nu),\ \mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right), \\ &\left(\frac{\nu}{\mu}\right)^s \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right), \\ &\mu \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \cdot \left(\frac{\mu}{\nu} - 1\right)\Bigg\}\end{aligned} \tag{11.2.10}$$

and with finite exception set

$$F = \{(n_1, n_2, k) : n_1 + n_2 = 0\}.$$

We define

$$\begin{aligned}\varepsilon = \min\Bigg\{&C \cdot (\mu - \nu),\ -\lambda_1 - p\,\lambda_2 + \mu - C \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right), \\ &-\lambda_1 - \lambda_2 + \mu - C \cdot \left(\frac{\nu}{\mu}\right)^s \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right), \\ &-\lambda_1 - \lambda_2 + \mu - C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \cdot \left(\frac{\mu}{\nu} - 1\right)\Bigg\}.\end{aligned}$$

Due to the choice of C in (11.2.9) and $\widetilde{C}$ in (11.2.10) it holds $\varepsilon > 0$.

► First, we will check $(\mathbf{Q} \cdot \mathcal{L})(n_1, n_2, k) < \infty$ for $(n_1, n_2, k) \in F$.
Since $0 < \lambda_1 < \infty$, $0 < \lambda_2 < \infty$ and $0 < \nu < \infty$,
for $k = 0$ holds

$$(\mathbf{Q} \cdot \mathcal{L})(0, 0, 0) = \nu \cdot (\mathcal{L}(0, 0, 1) - \mathcal{L}(0, 0, 0)) < \infty,$$

for $k = 1, \ldots, s$ holds

$$\begin{aligned}(\mathbf{Q} \cdot \mathcal{L})(0, 0, k) = {}& \lambda_1 \cdot (\mathcal{L}(1, 0, k) - \mathcal{L}(0, 0, k)) + p\,\lambda_2 \cdot (\mathcal{L}(0, 1, k) - \mathcal{L}(0, 0, k)) \\ &+ \nu \cdot (\mathcal{L}(0, 0, k+1) - \mathcal{L}(0, 0, k)) < \infty,\end{aligned}$$

for $k = s+1, \ldots, b-1$ holds

$$\begin{aligned}(\mathbf{Q} \cdot \mathcal{L})\,(0,0,k) = \lambda_1 \cdot (\mathcal{L}(1,0,k) - \mathcal{L}(0,0,k)) + \lambda_2 \cdot (\mathcal{L}(0,1,k) - \mathcal{L}(0,0,k)) \\ +\nu \cdot (\mathcal{L}(0,0,k+1) - \mathcal{L}(0,0,k)) < \infty,\end{aligned}$$

for $k = b$ holds

$$(\mathbf{Q} \cdot \mathcal{L})\,(0,0,b) = \lambda_1 \cdot (\mathcal{L}(1,0,b) - \mathcal{L}(0,0,b)) + \lambda_2 \cdot (\mathcal{L}(0,1,b) - \mathcal{L}(0,0,b)) < \infty.$$

► Second, we will check $(\mathbf{Q} \cdot \mathcal{L})\,(n_1, n_2, k) \le -\varepsilon$ for $z = (n_1, n_2, k) \notin F$ with

$$\begin{aligned}-\varepsilon = \max\Bigg\{ & C \cdot (\nu - \mu)\,,\ \lambda_1 + p\,\lambda_2 - \mu + C \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right), \\ & \lambda_1 + \lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^s \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right), \\ & \lambda_1 + \lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \cdot \left(\frac{\mu}{\nu} - 1\right)\Bigg\}.\end{aligned}$$

For $k = 0$ holds
for $n_1 = 1, 2, \ldots$ and $n_2 = 0, 1, 2, \ldots$

$$\begin{aligned}(\mathbf{Q} \cdot \mathcal{L})\,(n_1, n_2, 0) &= \nu \cdot (\mathcal{L}(n_1, n_2, 1) - \mathcal{L}(n_1, n_2, 0)) \\ &\overset{(11.2.7)}{=} \nu \cdot (n_1 + n_2 + \alpha(1) - n_1 - n_2 - \alpha(0)) \\ &\overset{(11.2.8)}{=} \nu \cdot \left(C - C \cdot \frac{\mu}{\nu}\right) = C \cdot \nu \cdot \left(1 - \frac{\mu}{\nu}\right) = C \cdot (\nu - \mu) \le -\varepsilon\end{aligned}$$

and for $n_1 = 0$ and $n_2 = 1, 2, \ldots$

$$\begin{aligned}(\mathbf{Q} \cdot \mathcal{L})\,(0, n_2, 0) &= \nu \cdot (\mathcal{L}(0, n_2, 1) - \mathcal{L}(0, n_2, 0)) \\ &\overset{(11.2.7)}{=} \nu \cdot (n_2 + \alpha(1) - n_2 - \alpha(0)) \\ &\overset{(11.2.8)}{=} \nu \cdot \left(C - C \cdot \frac{\mu}{\nu}\right) = C \cdot \nu \cdot \left(1 - \frac{\mu}{\nu}\right) = C \cdot (\nu - \mu) \le -\varepsilon.\end{aligned}$$

For $k = 1, \dots, s$ holds
for $n_1 = 1, 2, \dots$ and $n_2 = 0, 1, 2, \dots$

$$\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(n_1, n_2, k) \\
= \; & \lambda_1 \cdot (\mathcal{L}(n_1 + 1, n_2, k) - \mathcal{L}(n_1, n_2, k)) \\
& + p\,\lambda_2 \cdot (\mathcal{L}(n_1, n_2 + 1, k) - \mathcal{L}(n_1, n_2, k)) \\
& + \mu \cdot (\mathcal{L}(n_1 - 1, n_2, k - 1) - \mathcal{L}(n_1, n_2, k)) \\
& + \nu \cdot (\mathcal{L}(n_1, n_2, k + 1) - \mathcal{L}(n_1, n_2, k)) \\
\overset{(11.2.7)}{=} \; & \lambda_1 \cdot (n_1 + 1 + n_2 + \alpha(k) - n_1 - n_2 - \alpha(k)) \\
& + p\,\lambda_2 \cdot (n_1 + n_2 + 1 + \alpha(k) - n_1 - n_2 - \alpha(k)) \\
& + \mu \cdot (n_1 - 1 + n_2 + \alpha(k - 1) - n_1 - n_2 - \alpha(k)) \\
& + \nu \cdot (n_1 + n_2 + \alpha(k + 1) - n_1 - n_2 - \alpha(k)) \\
= \; & \lambda_1 + p\,\lambda_2 - \mu \\
& + \mu \cdot (\alpha(k - 1) - \alpha(k)) + \nu \cdot (\alpha(k + 1) - \alpha(k)) \\
\overset{(11.2.8)}{=} \; & \lambda_1 + p\,\lambda_2 - \mu \\
& + \mu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k-2} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) + \nu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) \\
= \; & \lambda_1 + p\,\lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \left(\frac{\mu}{\nu} - 1\right) + C \cdot \nu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\frac{\nu}{\mu} - 1\right) \\
= \; & \lambda_1 + p\,\lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right) \right) \le -\varepsilon
\end{aligned}$$

and for $n_1 = 0$ and $n_2 = 1, 2, \dots$

$$\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(0, n_2, k) \\
= \; & \lambda_1 \cdot (\mathcal{L}(1, n_2, k) - \mathcal{L}(0, n_2, k)) \\
& + p\,\lambda_2 \cdot (\mathcal{L}(0, n_2 + 1, k) - \mathcal{L}(0, n_2, k)) \\
& + \mu \cdot (\mathcal{L}(0, n_2 - 1, k - 1) - \mathcal{L}(0, n_2, k)) \\
& + \nu \cdot (\mathcal{L}(0, n_2, k + 1) - \mathcal{L}(0, n_2, k)) \\
\overset{(11.2.7)}{=} \; & \lambda_1 \cdot (1 + n_2 + \alpha(k) - n_2 - \alpha(k)) \\
& + p\,\lambda_2 \cdot (n_2 + 1 + \alpha(k) - n_2 - \alpha(k)) \\
& + \mu \cdot (n_2 - 1 + \alpha(k - 1) - n_2 - \alpha(k)) \\
& + \nu \cdot (n_2 + \alpha(k + 1) - n_2 - \alpha(k)) \\
\overset{(11.2.8)}{=} \; & \lambda_1 + p\,\lambda_2 - \mu \\
& + \mu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k-2} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) + \nu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) \\
= \; & \lambda_1 + p\,\lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\frac{\mu}{\nu} - 1\right) + C \cdot \nu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\frac{\nu}{\mu} - 1\right) \\
= \; & \lambda_1 + p\,\lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right) \right) \le -\varepsilon.
\end{aligned}$$

For $k = s+1, \ldots, b-1$ holds
for $n_1 = 1, 2, \ldots$ and $n_2 = 0, 1, 2, \ldots$

$$\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(n_1, n_2, k) \\
= \; & \lambda_1 \cdot (\mathcal{L}(n_1+1, n_2, k) - \mathcal{L}(n_1, n_2, k)) \\
& + \lambda_2 \cdot (\mathcal{L}(n_1, n_2+1, k) - \mathcal{L}(n_1, n_2, k)) \\
& + \mu \cdot (\mathcal{L}(n_1-1, n_2, k-1) - \mathcal{L}(n_1, n_2, k)) \\
& + \nu \cdot (\mathcal{L}(n_1, n_2, k+1) - \mathcal{L}(n_1, n_2, k)) \\
\stackrel{(11.2.7)}{=} \; & \lambda_1 \cdot (n_1 + 1 + n_2 + \alpha(k) - n_1 - n_2 - \alpha(k)) \\
& + \lambda_2 \cdot (n_1 + n_2 + 1 + \alpha(k) - n_1 - n_2 - \alpha(k)) \\
& + \mu \cdot (n_1 - 1 + n_2 + \alpha(k-1) - n_1 - n_2 - \alpha(k)) \\
& + \nu \cdot (n_1 + n_2 + \alpha(k+1) - n_1 - n_2 - \alpha(k)) \\
= \; & \lambda_1 + \lambda_2 - \mu \\
& + \mu \cdot (\alpha(k-1) - \alpha(k)) + \nu \cdot (\alpha(k+1) - \alpha(k)) \\
\stackrel{(11.2.8)}{=} \; & \lambda_1 + \lambda_2 - \mu \\
& + \mu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k-2} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) + \nu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) \\
= \; & \lambda_1 + \lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \left(\frac{\mu}{\nu} - 1\right) + C \cdot \nu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\frac{\nu}{\mu} - 1\right) \\
= \; & \lambda_1 + \lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right) \leq -\varepsilon
\end{aligned}$$

and for $n_1 = 0$ and $n_2 = 1, 2, \ldots$

$$\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(0, n_2, k) \\
= \; & \lambda_1 \cdot (\mathcal{L}(1, n_2, k) - \mathcal{L}(0, n_2, k)) \\
& + \lambda_2 \cdot (\mathcal{L}(0, n_2+1, k) - \mathcal{L}(0, n_2, k)) \\
& + \mu \cdot (\mathcal{L}(0, n_2-1, k-1) - \mathcal{L}(0, n_2, k)) \\
& + \nu \cdot (\mathcal{L}(0, n_2, k+1) - \mathcal{L}(0, n_2, k)) \\
\stackrel{(11.2.7)}{=} \; & \lambda_1 \cdot (1 + n_2 + \alpha(k) - n_2 - \alpha(k)) \\
& + \lambda_2 \cdot (n_2 + 1 + \alpha(k) - n_2 - \alpha(k)) \\
& + \mu \cdot (n_2 - 1 + \alpha(k-1) - n_2 - \alpha(k)) \\
& + \nu \cdot (n_2 + \alpha(k+1) - n_2 - \alpha(k)) \\
\stackrel{(11.2.8)}{=} \; & \lambda_1 + \lambda_2 - \mu \\
& + \mu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k-2} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) + \nu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{k} - C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \right) \\
= \; & \lambda_1 + \lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\frac{\mu}{\nu} - 1\right) + C \cdot \nu \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\frac{\nu}{\mu} - 1\right) \\
= \; & \lambda_1 + \lambda_2 - \mu + C \cdot \left(\frac{\nu}{\mu}\right)^{k-1} \cdot \left(\mu \cdot \left(\frac{\mu}{\nu} - 1\right) + \nu \cdot \left(\frac{\nu}{\mu} - 1\right)\right) \leq -\varepsilon.
\end{aligned}$$

For $k = b$ holds
for $n_1 = 1, 2, \dots$ and $n_2 = 0, 1, 2, \dots$

$$\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(n_1, n_2, b) \\
= \; & \lambda_1 \cdot (\mathcal{L}(n_1 + 1, n_2, b) - \mathcal{L}(n_1, n_2, b)) \\
& + \lambda_2 \cdot (\mathcal{L}(n_1, n_2 + 1, b) - \mathcal{L}(n_1, n_2, b)) \\
& + \mu \cdot (\mathcal{L}(n_1 - 1, n_2, b - 1) - \mathcal{L}(n_1, n_2, b)) \\
\overset{(11.2.7)}{=} \; & \lambda_1 \cdot (n_1 + 1 + n_2 + \alpha(b) - n_1 - n_2 - \alpha(b)) \\
& + \lambda_2 \cdot (n_1 + n_2 + 1 + \alpha(b) - n_1 - n_2 - \alpha(b)) \\
& + \mu \cdot (n_1 - 1 + n_2 + \alpha(b - 1) - n_1 - n_2 - \alpha(b)) \\
= \; & \lambda_1 + \lambda_2 - \mu + \mu \cdot (\alpha(b - 1) - \alpha(b)) \\
\overset{(11.2.8)}{=} \; & \lambda_1 + \lambda_2 - \mu + \mu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{b-2} - C \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \right) \\
= \; & \lambda_1 + \lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \left(\frac{\mu}{\nu} - 1\right) \leq -\varepsilon
\end{aligned}$$

and for $n_1 = 0$ and $n_2 = 1, 2, \dots$

$$\begin{aligned}
& (\mathbf{Q} \cdot \mathcal{L})(0, n_2, b) \\
= \; & \lambda_1 \cdot (\mathcal{L}(1, n_2, b) - \mathcal{L}(0, n_2, b)) \\
& + \lambda_2 \cdot (\mathcal{L}(0, n_2 + 1, b) - \mathcal{L}(0, n_2, b)) \\
& + \mu \cdot (\mathcal{L}(0, n_2 - 1, b - 1) - \mathcal{L}(0, n_2, b)) \\
\overset{(11.2.7)}{=} \; & \lambda_1 \cdot (1 + n_2 + \alpha(b) - n_2 - \alpha(b)) \\
& + \lambda_2 \cdot (n_2 + 1 + \alpha(b) - n_2 - \alpha(b)) \\
& + \mu \cdot (n_2 - 1 + \alpha(b - 1) - n_2 - \alpha(b)) \\
= \; & \lambda_1 + \lambda_2 - \mu + \mu \cdot (\alpha(b - 1) - \alpha(b)) \\
\overset{(11.2.8)}{=} \; & \lambda_1 + \lambda_2 - \mu + \mu \cdot \left(C \cdot \left(\frac{\nu}{\mu}\right)^{b-2} - C \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \right) \\
= \; & \lambda_1 + \lambda_2 - \mu + C \cdot \mu \cdot \left(\frac{\nu}{\mu}\right)^{b-1} \cdot \left(\frac{\mu}{\nu} - 1\right) \leq -\varepsilon.
\end{aligned}$$

□

11.3. Pure inventory system

In this section, we consider the case of zero service time, which is the version of our model in the classical inventory theory. The supply chain of interest is depicted in Figure 11.3.1 and consists of priority and ordinary customers, an inventory and a supplier.

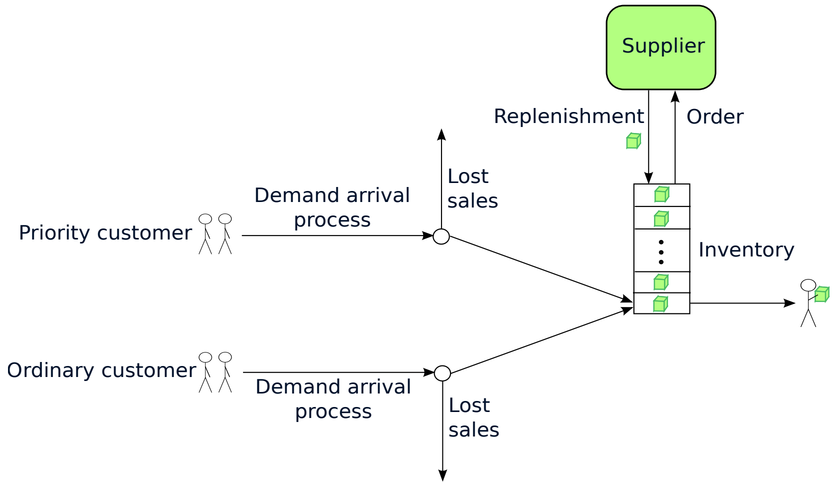


Figure 11.3.1.: The pure inventory system with two customer classes

There are two types of customers — priority customers and ordinary customers. The set of these customer classes is $\overline{C} = \{1, 2\}$, where 1 is the type of priority customers and 2 is the type of ordinary customers — the smaller the number, the higher the priority. According to two independent Poisson processes with different parameters the demands of each type of customer arrive one by one at the production system and require service. Priority customers arrive according to a Poisson process with rate $\lambda_1 > 0$ and ordinary customers arrive according to a Poisson process with rate $\lambda_2 > 0$.

Customers' arrivals are regulated by a flexible admission control with priority parameter p, $0 \leq p \leq 1$: If the inventory is depleted all arriving customers are rejected ("lost sales"). If the on-hand inventory is greater than a prescribed threshold level s, $0 < s < b$, the customers of both classes are admitted to enter the system. If the on-hand inventory reaches or falls below the threshold level s, all priority customers still enter the system, but ordinary customers are allowed to enter only with probability p and are rejected with probability $1 - p$.

The inventory is depleted by an exogenous customer demand and each customer needs exactly one item from the inventory.

It is assumed that transmission times for orders are zero and and that the transportation time between the production system and the inventory is negligible.

All inter-arrival times and replenishment lead times constitute an independent family of random variables.

An outside supplier replenishes raw material to the inventory according to the base stock policy. The replenishment lead time is exponentially distributed with parameter $\nu > 0$.

Let $Y(t)$ denote the on-hand inventory at time $t \geq 0$. Denote by $Y = (Y(t) : t \geq 0)$ the pure inventory process. Then, due to the usual independence and memoryless assumptions Y is a homogeneous strong Markov process. The state space of Y is

$$K = \{0, \dots, b\},$$

where b is the maximal size of the inventory, which depends on the replenishment policy.

The queueing-inventory process Y is irreducible. This can be shown analogously as in Appendix E on page 375 for the queueing-inventory system with (r, Q)-policy. From $|K| < \infty$ follows ergodicity (cf. [Ser13, Theorem 4.21]).

Definition 11.3.1. For the queueing-inventory process Z in a state space K, whose limiting distribution exists, we define

$$\theta := (\theta(k) : k \in K), \quad \theta(k) := \lim_{t \to \infty} P(Y(t) = k).$$

Proposition 11.3.2. *The inventory process $Y = (Y(t) : t \geq 0)$ has the following limiting and stationary distribution*

$$\begin{aligned}
\theta(0) &= \left[\sum_{j=0}^{s} \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^j + \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^s \cdot \sum_{j=1}^{b-s} \left(\frac{\nu}{\lambda_1 + \lambda_2} \right)^j \right]^{-1} \\
&= \left[\frac{1 - \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^{s+1}}{1 - \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)} + \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^s \cdot \left(\frac{1 - \left(\frac{\nu}{\lambda_1 + \lambda_2} \right)^{b-s+1}}{1 - \left(\frac{\nu}{\lambda_1 + \lambda_2} \right)} - 1 \right) \right]^{-1} \\
&= \left[\frac{1 - \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^{s+1}}{1 - \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)} + \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^s \cdot \left(\frac{\nu}{\lambda_1 + \lambda_2} \right) \cdot \left(\frac{1 - \left(\frac{\nu}{\lambda_1 + \lambda_2} \right)^{b-s}}{1 - \left(\frac{\nu}{\lambda_1 + \lambda_2} \right)} \right) \right]^{-1}, \\
\theta(k) &= \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^k \cdot \theta(0), \qquad k = 1, \dots, s, \\
\theta(k) &= \left(\frac{\nu}{\lambda_1 + p\,\lambda_2} \right)^s \left(\frac{\nu}{\lambda_1 + \lambda_2} \right)^{k-s} \cdot \theta(0), \qquad k = s+1, \dots, b.
\end{aligned}$$

Proof. The global balance equations of the pure inventory system with $s < b$ are given as follows

$$\begin{aligned}
&\theta(0) \cdot \nu = \theta(1) \cdot (\lambda_1 + p\,\lambda_2), \\
&\theta(k) \cdot (\lambda_1 + p\,\lambda_2 + \nu) = \theta(k+1) \cdot (\lambda_1 + p\,\lambda_2) + \theta(k-1) \cdot \nu, \quad k = 1, \dots, s-1, \\
&\theta(s) \cdot (\lambda_1 + p\,\lambda_2 + \nu) = \theta(s+1) \cdot (\lambda_1 + \lambda_2) + \theta(s-1) \cdot \nu, \\
&\theta(k) \cdot (\lambda_1 + \lambda_2 + \nu) = \theta(k+1) \cdot (\lambda_1 + \lambda_2) + \theta(k-1) \cdot \nu, \quad k = s+1, \dots, b-1, \\
&\theta(b) \cdot (\lambda_1 + \lambda_2) = \theta(b-1) \cdot \nu.
\end{aligned}$$

θ is the distribution of a finite birth-and-death process with birth-rates ν and death-rates $\lambda_1 + p\,\lambda_2$ for $k = 1, \dots, s$ and $\lambda_1 + \lambda_2$ for $k = s+1, \dots, b$ (cf. [Asm03, Corollary 2.5, p. 74]). $\square$

Remark 11.3.3. For $p = 1$, the model corresponds to the classic inventory system with base stock policy and arrival rate $\lambda_1 + \lambda_2$ and lost sales.

For $\lambda_2 \to 0$, the model corresponds to the classic inventory system with base stock policy and arrival rate λ_1.

11.4. Cost analysis

We consider the following cost structure for inventory, production and replenishment. The total costs consist of shortage costs $c_{ls,1}$ resp. $c_{ls,2}$ for each priority resp. ordinary customer that is lost, waiting costs $c_{w,1}$ resp. $c_{w,2}$ per unit of time for each priority resp. ordinary customer in the system (waiting or in service), capacity costs c_s per unit of time for providing inventory storage space (e.g. rent, insurance), holding costs c_h per unit of time for each unit that is kept on inventory. We assume that all of these costs per unit of time are positive.

Therefore, the cost function per unit of time in the respective states is

$$f_{b,p} : \ \mathbb{N}_0 \times \mathbb{N}_0 \times \{0, 1, \dots, b\} \quad \longrightarrow \quad \mathbb{R}_0^+$$

with

$$\begin{aligned} f_{b,p}(n_1, n_2, k) = {} & c_{w,1} \cdot n_1 + c_{w,2} \cdot n_2 + c_s \cdot b + c_h \cdot k + c_{ls,1} \cdot \lambda_1 \cdot 1_{\{k=0\}} \\ & + c_{ls,2} \cdot \left(\lambda_2 \cdot 1_{\{k=0\}} + (1-p) \cdot \lambda_2 \cdot 1_{\{0<k\le s\}} \right). \end{aligned}$$

The asymptotic average costs for an ergodic system can be computed similar to Section 10.3.

Appendix

A. Basics

A.1. Properties for Markov processes

In this section, we summarize some properties for Markov processes, which we need in some proofs.

Theorem A.1.1 (Cut criteria for Markov processes)**.**

(a) *Let Z be an irreducible and positive recurrent Markov process with state space E, infinitesimal generator* $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$, and stationary distribution $\pi := (\pi(z) : (z) \in E)$*. Then, the probability flux each way across a cut balances. That is for any $A \subset E$,*

$$\sum_{z\in A}\sum_{\tilde{z}\in A^c} \pi(z)\cdot q(z,\tilde{z}) = \sum_{z\in A}\sum_{\tilde{z}\in A^c} \pi(\tilde{z})\cdot q(\tilde{z},z).$$

(b) *Let Z be an irreducible and recurrent Markov process with state space E, infinitesimal generator* $\mathbf{Q} = (q(z;\tilde{z}) : z, \tilde{z} \in E)$, and *stationary measure* $\mathbf{x} = (x(z) : z \in E)$*. The probability flux each way across a finitely generated cut balances. That is for any finite $A \subset E$ (or $A \subset E$ with finite complement A^c),*

$$\sum_{z\in A}\sum_{\tilde{z}\in A^c} x(z)\cdot q(z,\tilde{z}) = \sum_{z\in A}\sum_{\tilde{z}\in A^c} x(\tilde{z})\cdot q(\tilde{z},z).$$

Proof.

(a) The proof for a positive recurrent Markov process is presented in [Kel79, Lemma 1.4, p. 8].

(b) If Z is irreducible and recurrent, there exists one, and up to a multiplicative factor only one, stationary measure $\mathbf{x} = (x(z) : z \in E)$. This x has the property $x(z) > 0$ for all $z \in E$ and can be found as solution of the balance equation $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ [Asm03, Theorem 4.2, p. 51].
The balance equation $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ can be written as

$$\begin{aligned} & \sum_{\tilde{z}\in E} x(z)\cdot q(z,\tilde{z}) = 0, \quad z \in E, \\ \Leftrightarrow \quad & x(z)\cdot(-q(z,z)) = \sum_{\tilde{z}\in E\setminus\{z\}} x(\tilde{z})\cdot q(\tilde{z},z), \quad z\in E, \\ \Leftrightarrow \quad & \sum_{\tilde{z}\in E\setminus\{z\}} x(z)\cdot q(z,\tilde{z}) = \sum_{\tilde{z}\in E\setminus\{z\}} x(\tilde{z})\cdot q(\tilde{z},z), \quad z\in E, \end{aligned}$$

because $-q(z,z) = \sum_{\tilde{z}\in E\setminus\{z\}} q(z,\tilde{z})$.
Summing up these balance equations over $z \in A$ yields

$$\sum_{z\in A}\sum_{\tilde{z}\in E\setminus\{z\}} x(z)\cdot q(z,\tilde{z}) = \sum_{z\in A}\sum_{\tilde{z}\in E\setminus\{z\}} x(\tilde{z})\cdot q(\tilde{z},z) < \infty,$$

since A is finite, which is equivalent to

$$\begin{aligned}&\sum_{z\in A}\sum_{\tilde{z}\in A\setminus\{z\}} x(z)\cdot q(z,\tilde{z}) + \sum_{z\in A}\sum_{\tilde{z}\in A^c} x(z)\cdot q(z,\tilde{z})\\ =&\sum_{z\in A}\sum_{\tilde{z}\in A\setminus\{z\}} x(\tilde{z})\cdot q(\tilde{z},z) + \sum_{z\in A}\sum_{\tilde{z}\in A^c} x(\tilde{z})\cdot q(\tilde{z},z).\end{aligned}$$

Hence,

$$\sum_{z\in A}\sum_{\tilde{z}\in A^c} x(z)\cdot q(z,\tilde{z}) = \sum_{z\in A}\sum_{\tilde{z}\in A^c} x(\tilde{z})\cdot q(\tilde{z},z).$$

□

Foster [Fos53] introduced a technique for proving stability of Markov chains. Many variants of the Foster-Lyapunov stability criterion exist. The following version for continuous time Markov processes has been adapted from Kelly and Yudovina [KY14, Proposition D.3].

Theorem A.1.2 (Foster-Lyapunov stability criterion for Markov processes)**.** *Let Z be a (time-homogeneous, irreducible, non-explosive, conservative) continuous time Markov process with countable state space E and matrix of transition rates $\mathbf{Q} = (q(z;\tilde{z}) : z,\tilde{z} \in E)$. Suppose $\mathcal{L} : E \to \mathbb{R}_0^+$ is a function such that, for some constants $\varepsilon > 0$, some finite exception set $F \subsetneq E$, and all $z \in E$*

$$(\mathbf{Q}\cdot\mathcal{L})(z) = \sum_{\tilde{z}\in E} q(z;\tilde{z})\ (\mathcal{L}(\tilde{z}) - \mathcal{L}(z)) \begin{cases}\leq -\varepsilon, & z \notin F,\\ < \infty, & z \in F.\end{cases} \tag{A.1.1}$$

Then the expected return time to K is finite, and Z is positive recurrent.[1]

[1] In [KY14, Proposition D.3] it is $\sum_{\tilde{z}\in E} q(z;\tilde{z})\ (\mathcal{L}(\tilde{z}) - \mathcal{L}(z)) < b - \varepsilon,\ z \in F$, for some b.

A.2. Standard separable networks

In this section, we summarize definitions and theorems on classical exponential networks. The open queueing network has been introduced by Jackson [Jac57]. The closed analogue of the Jackson network — the Gordon-Newell network — has been introduced some years later by Jackson, see [Jac63], and has been rediscovered by Gordon and Newell [GN67].

We sketch only the relevant results for queueing networks of single server nodes with state dependent service intensities for our research. The results for a Jackson network with multi-server stations is summarized in [Dad01b, Definition 2.5, Theorem 2.6, pp. 313f.].

The notation in the next definition follows [SD03, Definition 2.1, pp. 168].

Definition A.2.1 (Jackson network)**.** A Jackson network is a network of J numbered service stations (nodes), denoted by $\overline{J} := \{1, 2, \ldots, J\}$. Each station j consists of a single server with infinite waiting room under FCFS regime. Customers in the network are indistinguishable. There is an external Poisson-λ_j arrival stream at node j with $\lambda_j \geq 0$. Customers arrive at node j from the outside or from other nodes of the network and request for a service there. The service time is exponentially distributed with mean 1. If there are $n_j > 0$ customers present at node j, service at node j is provided with intensity $\mu_j(n_j) > 0$, and otherwise $\mu_j(0) := 0$. All service and inter-arrival times constitute an independent family of random variables.

Movements of the customers in the network are governed by a Markovian routing mechanism: A customer when leaving node i selects with probability $r(i,j) \geq 0$ to visit node j next, and then enters node j immediately, starting service if he finds the server idle, otherwise he joins the tail of the queue at node j; this customer decides to leave the network immediately with probability $r(i,0) \geq 0$ ($\sum_{j=0}^{J} r(i,j) = 1$ holds for all $i \in \overline{J}$). The artificial node 0 represents the external source and sink of the network. Given the departure node i the customer's routing decision is made independently of the network's history.

We assume that with $\overline{J}_0 := \{0, 1, \ldots, J\}$, $\lambda := \sum_{j=1}^{J} \lambda_j$, $r(0,j) := \frac{\lambda_j}{\lambda}$ and $r(0,0) := 0$ the matrix $\mathcal{R} := \big(r(i,j) : i,j \in \overline{J}_0\big)$ is irreducible.

Let $X_j(t)$ denote the number of customers present at node j at time $t \geq 0$, either waiting or in service (local queue length at node j). Then $X(t) := (X_j(t) : j = 1, \ldots, J)$ is the local queue length vector of the network at time $t \geq 0$. We denote by $X = (X(t) : t \geq 0)$ the joint queue length process of the Jackson network with state space $E := \mathbb{N}_0^J$.

Theorem A.2.2 ([Jac57])**.** *The joint queue length process X of the Jackson network is a Markov process with the following transition rates $(q(x;y) : x, y \in E)$:*
For $i, j \in \overline{J}$, $i \neq j$, and $x = (n_1, \ldots, n_J) \in E$

$$q\left((n_1, \ldots, n_i, \ldots, n_J);(n_1, \ldots, n_i + 1, \ldots, n_J)\right) = \lambda_i,$$

$$q\left((n_1, \ldots, n_i, \ldots, n_J);(n_1, \ldots, n_i - 1, \ldots, n_J)\right) = \mu_i(n_i) \cdot r(i,0) \cdot 1_{\{n_i>0\}},$$

$$q\left((n_1, \ldots, n_i, \ldots, n_j, \ldots, n_J);(n_1, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_J)\right) = \mu_i(n_i) \cdot r(i,j) \cdot 1_{\{n_i>0\}}.$$

Furthermore,

$$q(x;x) = -\sum_{\substack{y \in E,\\ y \neq x}} q(x;y) \quad \text{and} \quad q(x;y) = 0 \text{ otherwise.}$$

The traffic equation of the network is defined by

$$\eta_j = \lambda_j + \sum_{i=1}^{J} \eta_i \cdot r(i,j), \quad j \in \overline{J},$$

and has a unique solution which we denote by $\eta = (\eta_1, \ldots, \eta_J)$.

We assume henceforth X *to be ergodic. The unique stationary and limiting distribution* π *of* X *is then*

$$\pi(n_1, \ldots, n_J) = \prod_{j \in \overline{J}} \pi_j(n_j) = \prod_{j \in \overline{J}} C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)}, \quad (n_1, \ldots, n_J) \in \mathbb{N}_0^J, \tag{A.2.1}$$

with finite normalisation constants C_j, $j \in \overline{J}$.

Remark A.2.3 (Product form). [Dad01b, Remark 2.7, p. 314] The stationary distribution (A.2.1) is a so-called product form distribution. This means that the joint stationary distribution π for the network process is a product of its marginal distributions π_j. Hence, in equilibrium the local queue lengths at a fixed time point behave as if they are independent. However, it should be stressed that the processes X_j, $j = 1, \ldots, J$, are by no means independent since the queue lengths rely on the customers moving between those nodes.

The form of the local equilibria π_j, $j \in \overline{J}$, suggests that in steady state node j behaves like an $M/M/1/\infty$-FCFS system in isolation with Poisson-η_j arrival streams and exponential-μ_j service times. However, it can be proven that in general the suggestion of Poisson streams between the nodes is not true (cf. [Mel79]).

Remark A.2.4 (Ergodicity). If the service intensities are independent of the queue length, i.e. $\mu_j(n_j) = \mu_j$ for all $n_j > 0$ and $j \in \overline{J}$, then the joint queue length process X is ergodic if and only if $\eta_j < \mu_j$ for all $j \in \overline{J}$.

The notation in the next definition follows [SD03, Definition 2.4, pp. 169f.].

Definition A.2.5 (Gordon-Newel network)**.** A Gordon-Newell network consists of a set of single server nodes $\overline{J} := \{1, 2, \ldots, J\}$ as described in Definition A.2.1 of the Jackson network without external arrivals and departures. This means that the probabilities $r(j, 0)$ for all $j \in \overline{J}$ and the total network arrival rate λ are set to zero. There are $D > 0$ indistinguishable customers cycling in the network according to an irreducible Markov matrix $\mathcal{R} = \big(r(i,j) : i, J \in \overline{J}\big)$. The service times are similar to the open Jackson network and the independence assumptions on service times and routing decisions are assumed to hold as well.

Let $X_j(t)$ denote the number of customers present at node j at time $t \geq 0$, either waiting or in service (local queue length at node j), then $X(t) := (X_j(t) : j = 1, \ldots, J)$ is the local queue length vector of the network at time $t \geq 0$.

We denote by $X = (X(t) : t \geq 0)$ the joint queue length process of the Gordon-Newell network with state space $S(J, D) := \big\{(n_1, n_2, \ldots, n_J) \in \mathbb{N}_0^J : n_1 + n_2 + \ldots + n_J = D\big\}$.

Theorem A.2.6 ([Jac63, GN67]). *The joint queue length process X of the Gordon-Newell network is a Markov process with the following transition rates $(q(x;y) : x, y \in S(J;D))$: For $i, j \in \overline{J}$ and $x = (n_1, \ldots, n_J) \in S(J,D)$*

$$q\left((n_1, \ldots, n_i, \ldots, n_j, \ldots, n_J); (n_1, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_J)\right) = \mu_i(n_i) \cdot r(i,j) \cdot 1_{\{n_i > 0\}}.$$

Furthermore,

$$q(x;x) = -\sum_{\substack{y \in E,\\ y \neq x}} q(x;y) \quad \text{and} \quad q(x;y) = 0 \text{ otherwise.}$$

The joint queue length process X is ergodic. Let $\eta = (\eta_1, \ldots, \eta_J)$ denote the unique probability solution of the traffic equation

$$\eta_j = \sum_{i=1}^{J} \eta_i \cdot r(i,j), \quad j \in \overline{J}.$$

The unique stationary and limiting distribution $\widehat{\pi} = \widehat{\pi}(J,D)$ of X on $S(J,D)$ is

$$\widehat{\pi}\,(n_1, \ldots, n_J) = G(J,D)^{-1} \prod_{\ell=1}^{J} \prod_{i=1}^{n_\ell} \left(\frac{\eta_\ell}{\mu_\ell(i)} \right), \quad (n_1, \ldots, n_J) \in S(J,D) \tag{A.2.2}$$

with $G(J,D)$ as normalisation constant.

Remark A.2.7. [Dad01b, Remark 2.11, p. 316] The stationary distribution $\pi(\cdot)$ of the Gordon-Newell network is said to be of product form as well, although the normalisation constant $G(J,D)$ does not factorise over the nodes as in the Jackson network. The stationary distribution (A.2.2) looks like being obtained by conditioning from an equilibrium in a Jackson network with the same nodes and suitably redefined routing to the set $\overline{J}$, given the number of customers present.

Baskett, Chandy, Muntz, and Palacios [BCMP75] and Kelly [Kel76] develop more complex product form models:

A **Kelly network** is a general multiclass Jackson network, which connects as set of quasi-reversible queues through some very general routing schemes. Instead of probabilistic routing, for each class of customers a fixed route through the network is defined. It may comprise symmetric servers with class-dependent generally distributed service times. A Kelly network still enjoys the basic properties of Jackson network, namely, the product form equilibrium distribution and the Poisson-in-Poisson-out property.

The results of Jackson and of Gordon and Newell have been extended in [BCMP75] by Baskett and his coauthors to queueing networks with several job classes, different queueing disciplines, and generally distributed service times for specific server types. The **BCMP networks** can be open, closed, or mixed (that contain open and closed classes). The BCMP theorem says that the BCMP networks have product form solution.

A.3. Jackson network in a random environment

As mentioned previously, comparing our production-inventory-replenishment systems in Chapter 2, Chapter 3, Section 4.3, Chapter 5, Chapter 6 and Chapter 7 with the "Jackson network in a random environment" in [KDO16, Section 4] it turns out that we can interpret the inventory-replenishment subsystem as a "random environment" for the production network of nodes $\overline{J}$, which is in this view a Jackson network of parallel servers.

In this section, we exemplary explain for our basic model in Chapter 2 the parameters "Jackson network in a random environment" of [KDO16, Section 4], so that for the interested reader the technique in [KDO16, Section 4] to control a Jackson network in a nonautonomous environment is easier comprehensible.

In our model, which is depicted in Figure A.3.1, the network process is $(X_1, \dots, X_J)$ with state space $\mathbb{N}^J$ and the environment process is $(Y_1, \dots, Y_J, W_{J+1})$ with state space K.

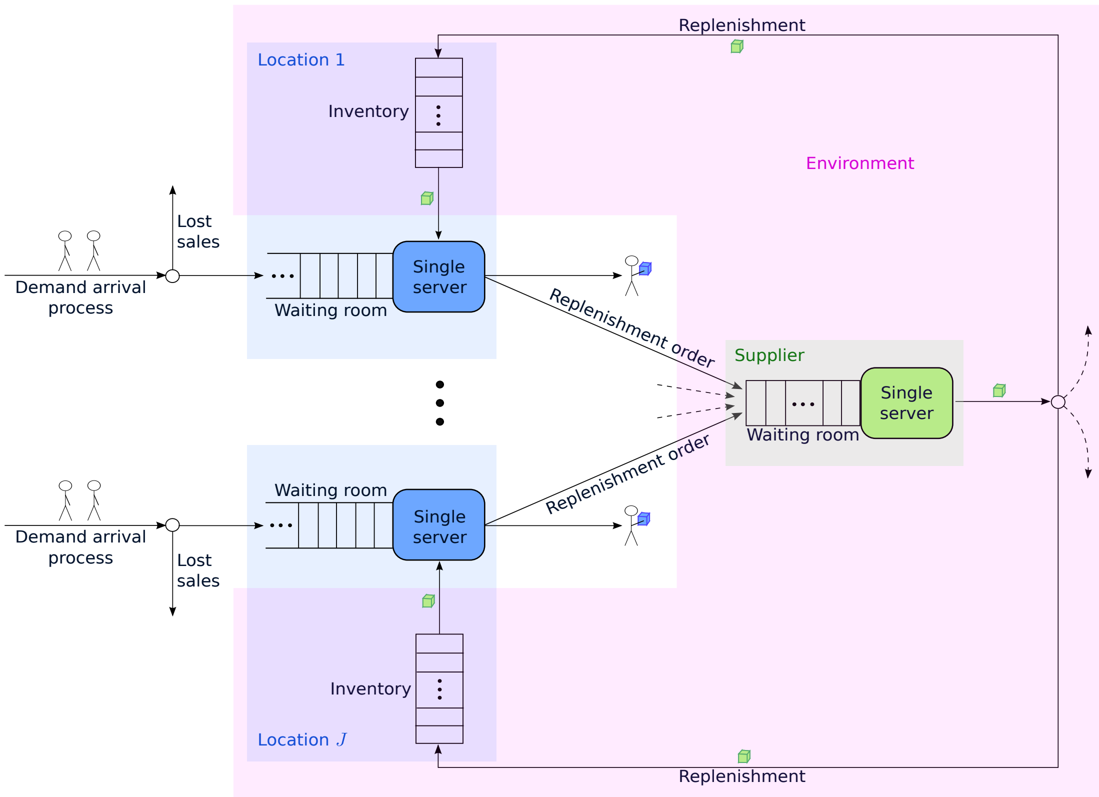


Figure A.3.1.: Supply chain with base stock policy

In the following, we will explain the dynamic of the interacting system.

Whenever the environment is in state $\mathbf{k} = (k_1, \ldots, k_J, k_{J+1}) \in K$ at time $t \geq 0$, then the environment changes its status to $\widetilde{\mathbf{k}} \in K$ with rate $v(\mathbf{k}, \widetilde{\mathbf{k}})$.

We set $V = \left(v(\mathbf{k}, \widetilde{\mathbf{k}}) : \mathbf{k}, \widetilde{\mathbf{k}} \in K\right)$. V is a generator matrix with

$$\begin{aligned} v(\mathbf{k}, \mathbf{k} + \mathbf{e}_i) &= \nu, && \text{if } 0 \leq k_i < b_i, \\ v(\mathbf{k}, \widetilde{\mathbf{k}}) &= 0, && \text{otherwise for } \mathbf{k} \neq \widetilde{\mathbf{k}}, \\ v(\mathbf{k}, \mathbf{k}) &= - \sum_{\widetilde{\mathbf{k}} \in K \setminus \{\mathbf{k}\}} v(\mathbf{k}, \widetilde{\mathbf{k}}). \end{aligned}$$

Whenever the environment is in state $\mathbf{k} \in K$ and at node $j \in \overline{J}$ a customer is served and leaves the network at time $t \geq 0$, then this jump of the local queue length triggers with probability $R_j(\mathbf{k}, \widetilde{\mathbf{k}})$ the environment to jump immediately from state $\mathbf{k}$ to state $\widetilde{\mathbf{k}} \in K$. We set $R_j = \left(R_j(\mathbf{k}, \widetilde{\mathbf{k}}) : \mathbf{k}, \widetilde{\mathbf{k}} \in K\right)$, $j \in \overline{J}$. R_j are stochastic matrices with

$$\begin{aligned} R_j(\mathbf{k}, \mathbf{k} - \mathbf{e}_j) &= 1, && \text{if } 1 \leq k_j \leq b_j, \\ R_j\,(\mathbf{k}, \mathbf{k}) &= 1, && \text{if } k_j = 0, \\ R_j(\mathbf{k}, \widetilde{\mathbf{k}}) &= 0, && \text{otherwise.} \end{aligned}$$

Associated with the environment state $\mathbf{k} \in K$ is a vector $\boldsymbol{\gamma}(\mathbf{k})$ which determines the factor by which the service capacities are changed, when the environment enters $\mathbf{k}$. We set $\boldsymbol{\gamma}(\mathbf{k}) = \left(\gamma_j(\mathbf{k}) : j \in \overline{J}\right)$ with

$$\gamma_j(\mathbf{k}) = \begin{cases} 0, & \text{if } k_j = 0 \quad \text{(i.e. service down)}, \\ 1, & \text{if } k_j \geq 1 \quad \text{(i.e. service up)}. \end{cases}$$

This results in state-dependent service rates $\mu_j(n_j, \mathbf{k}) = \gamma_j(\mathbf{k}) \cdot \mu_j(n_j)$ if the queue length at node j is n_j and the environment is in state $\mathbf{k}$.

In our system the arrival rate to the network is not increased as a reaction to the impact of the environment in state $\mathbf{k}$. Hence, $\beta(\mathbf{k}) = 1$.

In our model, the reaction to the servers' changes of capacities is as in randomized skipping as well as randomized reflection. We will explain the parameters for randomized skipping in more detail.

The routing matrix of customers (demand) is given by $r = \left(r(i,j) : i, j \in \overline{J}_0\right)$ with

$$\begin{aligned} r(0,j) &= \frac{\lambda_j}{\sum_{i \in \overline{J}} \lambda_i}, && j \in \overline{J}, \\ r(j,0) &= 1, && j \in \overline{J}, \\ r(i,j) &= 0, && \text{otherwise } i, j \in \overline{J}_0. \end{aligned}$$

$\overline{J}_0$ is an extended node set, where "0" refers to the external source and sink of the network.

The lost sales behaviour of customers (demand) can be represented by rerouting with either randomized skipping or randomized reflection. For the case of rerouting with randomized skipping, e.g. we need environment dependent rerouting with acceptance probabilities $\alpha = \alpha\left(\boldsymbol{\gamma}(\mathbf{k})\right) = \left(\alpha_j\left(\gamma_j(\mathbf{k}) : j \in \overline{J}_0\right)\right)$ with

$$\begin{aligned} \alpha_j(0) &= 0 \quad \text{(i.e. not accepted)}, \\ \alpha_j(1) &= 1 \quad \text{(i.e. accepted)}. \end{aligned}$$

Hence, the modified routing matrix is given by $r^{(\alpha(\boldsymbol{\gamma}(\mathbf{k})))} = \left(r^{(\alpha(\boldsymbol{\gamma}(\mathbf{k})))}(i,j) : i,j \in \overline{J}_0\right)$ with

$$\begin{aligned} r^{(\alpha(\boldsymbol{\gamma}(\mathbf{k})))}(0,0) &= \frac{\lambda_j}{\sum_{i \in \overline{J}^\star} \lambda_i}, && \overline{J}^\star := \left\{i \in \overline{J} : k_i = 0\right\}, \\ r^{(\alpha(\boldsymbol{\gamma}(\mathbf{k})))}(0,j) &= \frac{\lambda_j}{\sum_{i \in \overline{J}} \lambda_i}, && j \in \overline{J} \setminus \left\{i \in \overline{J} : k_i = 0\right\}, \\ r^{(\alpha(\boldsymbol{\gamma}(\mathbf{k})))}(j,0) &= 1, && j \in \overline{J}, \\ r^{(\alpha(\boldsymbol{\gamma}(\mathbf{k})))}(i,j) &= 0, && \text{otherwise } i,j \in \overline{J}_0. \end{aligned}$$

B. Appendix to Chapter 2

B.1. Iterative Algorithm

R code. The iterative algorithm on page 25 can be modeled in R. The R code is presented in the following.

Functions in R to calculate the average sojourn time $E[T]$ (for step 1 in the algorithm)

```
prob_tilde_x2=function(lambda_sup,nu,b){
 res<-(lambda_sup/nu)^(0:b)
 return(res/sum(res))
}


ET2<-function(lambda_sup,nu,b){
 prob_tilde_x2_tk<-prob_tilde_x2(lambda_sup,nu,b-1)
 if(b>=1){
  return(sum(prob_tilde_x2_tk*(1:b))/nu)
 }
 else{
  return(1/nu)
  }
 }
```

Functions in R to calculate the blocking probabilities q_j (for step 2 in the algorithm)

```
blockprob_1<-function(lambda_1,ET2,b_1){
 res1<-(((lambda_1/ET2)^(0:b_1))*(1/(factorial(0:b_1))))
norm.res1<-(res1/sum(res1))
return(tail(norm.res1,n=1))
}
#
blockprob_2<-function(lambda_2,ET2,b_2){
 res2<-(((lambda_2/ET2)^(0:b_2))*(1/(factorial(0:b_2))))
norm.res2<-(res2/sum(res2))
return(tail(norm.res2,n=1))
}
```

R-code to find the routing probabilities p_j

```
pfind<-function(lambda_1,lambda_2,b_1,b_2,nu){
 #Initialize
 b<-b_1+b_2          #
 tol<-0.001          # stop criterion
 q_1<-0              # blocking probability at location 1
 q_2<-0              # blocking probability at location 2
 q_1_old<-1-q_1+tol  # old blocking probability at location 1
 q_2_old<-1-q_2+tol  # old blocking probability at location 2
 steps<-0            # number of steps
 max.steps<-500      # maximal number of steps
 #
 while(abs(q_1-q_1_old)+abs(q_2-q_2_old)>tol&&steps<max.steps){
  # Step (1):
  # Calculation of effective arrival rates of replenishment orders
  # at the central supplier
  lambda_1_new<-(1-q_1)*lambda_1    #effective arrival rate from location 1
```

```
    lambda_2_new<-(1-q_2)*lambda_2   #effective arrival rate from location 2
    lambda_sup<-lambda_1_new+lambda_2_new
    #
    # Step (2):
    # Calcuation of average sojourn time of a replenishment order at the central supplier
    T_sup<-ET2(lambda_sup,nu,b)
    #
    # Step (3):
    # Determination of new blocking probabilities
    q_1_old<-q_1   #old blocking probability at location 1
    q_2_old<-q_2   #old blocking probability at location 1
    q_1<-blockprob_1(lambda_1_new,1/T_sup,b_1) #new blocking probability at location 1
    q_2<-blockprob_2(lambda_2_new,1/T_sup,b_2) #new blocking probability at location 2
    #
    # Check the stop criterion
    steps=steps+1  # counting steps
    #
    # if the algorithm does not converges in the maximal number of steps,
    # stop the algorithm
    if(steps>=max.steps){
    q_1<-NA
    }
  }
  #
  # Calculation of the routing probabilities
  p_1<-((1-q_1)*lambda_1)/((1-q_1)*lambda_1+(1-q_2)*lambda_2) #routing probability at location 1
  p_2<-((1-q_2)*lambda_2)/((1-q_1)*lambda_1+(1-q_2)*lambda_2) #routing probability at location 2
}
```

Functions in R to calculate γ_j

```
gamma1find<-function(lambda_1,greynu,p){
gamma1<-((greynu*p)/lambda_1) return(gamma1)
}
#
gamma2find<-function(lambda_2,greynu,p){
gamma2<-((greynu*(1-p))/lambda_2) return(gamma2)
}
```

Functions in R to calculate $P(Y_j = 0)$ (cf. (2.5.3) on page 31)

```
PY1eq0find<-function(gamma1,b_1){
if(gamma1!=1){
return((1-gamma1)/(1-gamma1^(b_1+1)))
}
else{
return(1/(b_1+1))
}
}
#
PY2eq0find<-function(gamma2,b_2){
if(gamma2!=1){
return((1-gamma2)/(1-gamma2^(b_2+1)))
}
else{
return(1/(b_2+1))
}
}
```

Functions in R to calculate $E(Y_j)$ (cf. (2.5.2) on page 31)

```
EY1find<-function(gamma1,b_1){
if(gamma1!=1){
return((gamma1/(1-gamma1))*((b_1*gamma1^(b_1+1)-(b_1+1)*gamma1^b_1 +1)/(1-
    gamma1^(b_1+1))))
}
else{
return(b_1/2)
}
}
#
EY2find<-function(gamma2,b_2){
if(gamma2!=1){
return((gamma2/(1-gamma2))*((b_2*gamma2^(b_2+1)-(b_2+1)*gamma2^b_2 +1)/(1-
    gamma2^(b_2+1))))
}
else{
return(b_2/2)
}
}
```

C. Appendix to Chapter 3

C.1. Algorithm to obtain $\widetilde{\theta}$

To obtain the exact or approximate solutions of queueing models various analytical, numerical and simulation techniques are available (e.g. the power iteration method, generating function approach, product form solution and recursive solution technique) (cf. [CK77, Section III, pp. 44ff.]). The recursive solution technique was first suggested by Herzog, Woo and Chandy [HWC75]. They demonstrate an efficient solution for single queueing models with other than exponential arrival or service time distributions. The recursive technique uses the fact that the steady state probabilities of the system can sometimes be expressed in terms of other steady state probabilities. This leads to a reduction of the number of unknowns in the global balance equations.

In this section, we assume that there are two heterogeneous locations with base stock levels $b_1 \geq b_2$, where $b_1 > 1$ and $b_2 \geq 1$ and arrival rates $\lambda_1, \lambda_2 > 0$. The state transition diagram for such a system is presented in Figure C.1.1. In Section 3.3.1.1 on page 50, we have derived an explicit solution for $\widetilde{\theta}(\mathbf{k})$, $\mathbf{k} \in K$, for the special case with base stock levels $b_1 = b_2 = 1$.

The steady state probabilities can be obtained by a recursive method which is described by the algorithm given below. This recursive solution technique uses the fact that the steady state probabilities of the system can be expressed in terms of other steady state probabilities. As mentioned above, this leads to a reduction of the number of unknowns in the global balance equations. κ is a variable which represents a temporarily unknown probability and GBE is used to denote a global balance equation.

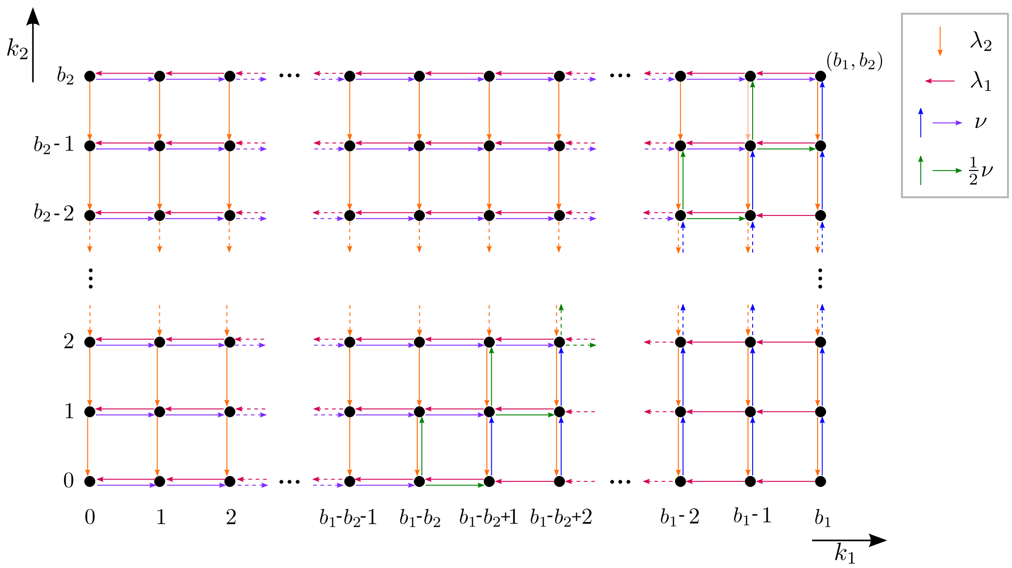


Figure C.1.1.: State transition diagram of a system with two heterogeneous locations

Chow and Kohler present in [CK79] a generalization of the recursive solution technique. They apply the method to non-homogeneous (= heterogeneous) two-processor systems with special properties in [CK79, Section IV, pp. 358f.] and present a sample system using the algorithm in [CK79, Appendix, pp. 360f.]. One special property of the load balancing policy that enables to use their recursive solution technique for the two-processor heterogeneous systems is that the policy line (= continuous chain of arrival transitions starting at state $(0,0)$ given that no departure occurs) partitions the states of the state transition rate diagram into two regions. Our load balancing policy is slightly different from that of Chow and Kohler [CK79, Section IV, pp. 358f.] because of $\frac{1}{2}\nu$ (since if both inventories have the same difference between the on-hand inventory and the capacity of the inventory, it enters either with equal probability) as can be seen in the state transition diagram in Figure C.1.1.

They mentioned that the generalization of recursive solution "technique for three or more processors does not appear to be straightforward" [CK79, p. 359].

We will henceforth use an abbreviated notation because $k_{J+1} = \sum_{j=1}^{J}(b_j - k_j)$ and the base stock levels b_j, $j \in \overline{J}$, are fixed parameters:

$$\widetilde{\theta}\Big(\overbrace{k_1,\quad k_2}^{\substack{\text{inventories}\\ \text{at locations}}}\Big) := \widetilde{\theta}\Big(\overbrace{k_1,\quad k_2,}^{\substack{\text{inventories}\\ \text{at locations}}}\ \overbrace{k_3}^{\text{supplier}}\Big)$$

and hence,

$$\begin{aligned} p_i(\mathbf{k} - \mathbf{e}_i) &:= p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}), \\ p_i(\mathbf{k} + \mathbf{e}_i) &:= p_i(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}), \quad i \in \overline{J}. \end{aligned}$$

In this section, we present the recursive method to obtain $\widetilde{\theta}(\mathbf{k})$, $\mathbf{k} \in K$, from the global balance equations $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$ (cf. equation (3.3.2) on page 47)

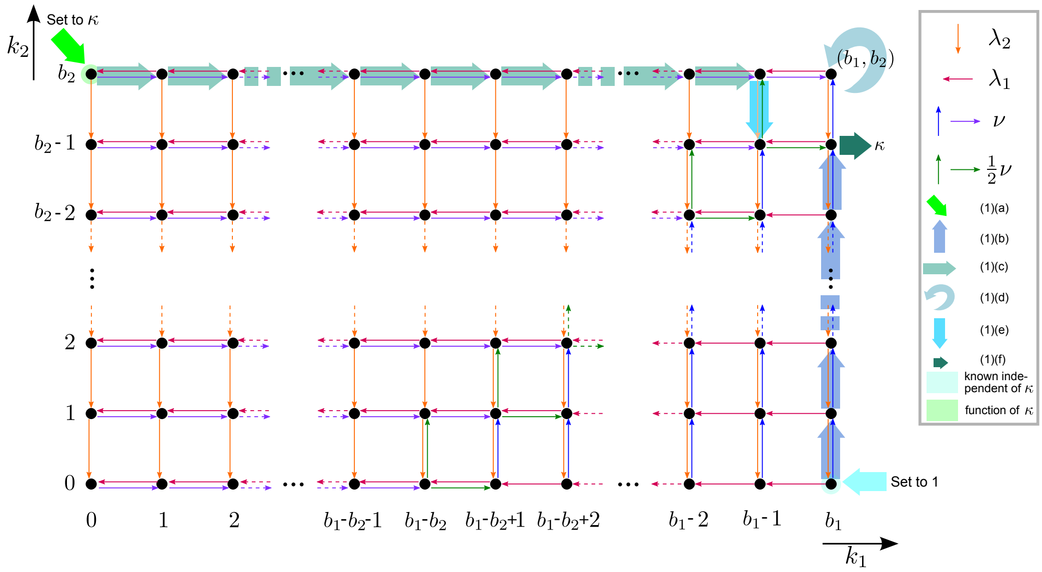
$$\begin{aligned} &\widetilde{\theta}\,(\mathbf{k}) \Big(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}} \Big) \\ = &\sum_{i \in \overline{J}} \widetilde{\theta}\,(\mathbf{k} + \mathbf{e}_i) \cdot \lambda_i \cdot 1_{\{k_i < b_i\}} \\ &+ \sum_{i \in \overline{J}} \widetilde{\theta}\,(\mathbf{k} - \mathbf{e}_i) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i) \cdot 1_{\{k_i > 0\}}, \end{aligned}$$

where $p_i(\mathbf{k})$, $i \in \overline{J}$, includes the cases

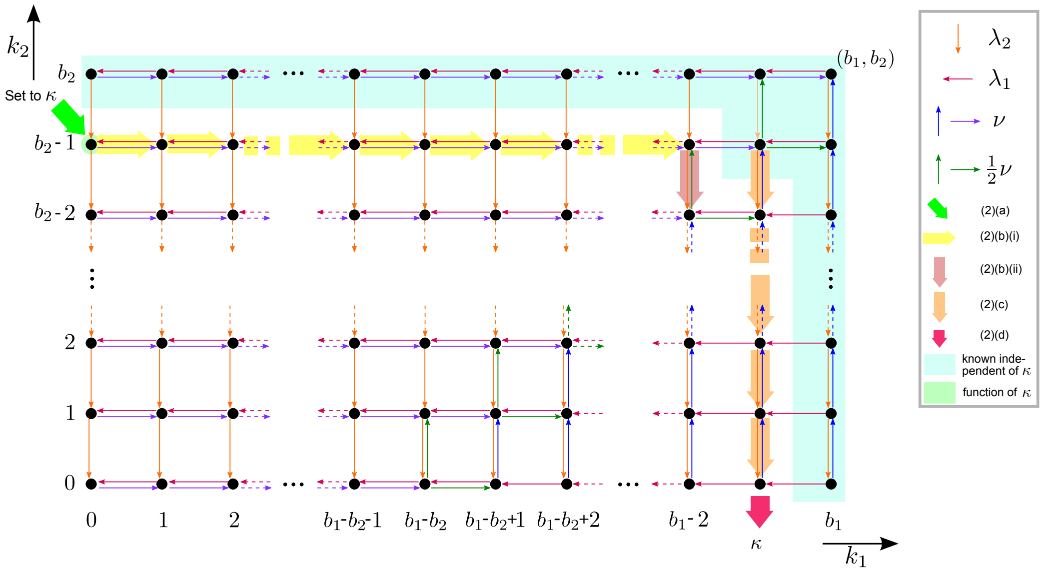
$$p_i(\mathbf{k}) = \begin{cases} 1, & \text{if } \{i\} = \arg\max\limits_{j \in \overline{J}}(b_j - k_j), \\ \frac{1}{|\arg\max\limits_{j \in \overline{J}}(b_j - k_j)|} < 1, & \text{if } \{i\} \subsetneq \arg\max\limits_{j \in \overline{J}}(b_j - k_j), \\ 0, & \text{if } i \notin \arg\max\limits_{j \in \overline{J}}(b_j - k_j). \end{cases}$$

On the next pages, we present an algorithm to obtain $\widetilde{\theta}(\mathbf{k})$ for a system with two heterogeneous locations and $b_1 \geq b_2$ and $b_1 > 1$, $b_2 > 1$. The algorithm for this system is illustrated by a simple example in Appendix C.1 on page 312. A few steps of the algorithm are visualised in the state transition diagram in Figure C.1.2.

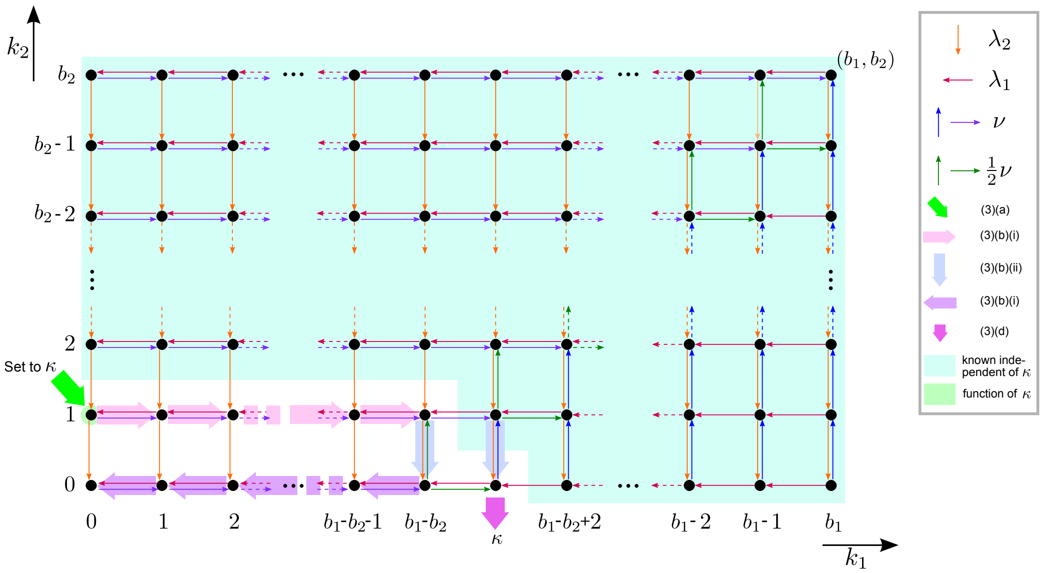
The algorithm for a system with base stock levels $b_1 > b_2 = 1$ is presented on page 320.



(a) First loop with $k_2 = b_2$



(b) Second loop with $k_2 = b_2 - 1$



(c) Last loop with $k_2 = 1$

Figure C.1.2.: Visualisation of a few steps of the algorithm

ALGORITHM ($b_1 \geq b_2$ with $b_1 > 1$, $b_2 > 1$ and $\lambda_1, \lambda_2 > 0$)

- Set $\widetilde{\theta}(b_1, 0) = 1$
- For $k_2 = b_2, \ldots, 1$

  (1) if $k_2 = b_2$,

  - **(a)** set $\widetilde{\theta}(0, b_2) = \kappa$
  - **(b)** for $\ell = 0, \ldots, b_2 - 2$
    use the GBE of state (b_1, ℓ)
    to find an expression for $\widetilde{\theta}(b_1, \ell + 1)$ independent of κ
  - **(c)** for $k_1 = 0, \ldots, b_1 - 2$
    use the GBE of state (k_1, b_2)
    to find an expression for $\widetilde{\theta}(k_1 + 1, b_2)$ as a function of κ
  - **(d)** use the GBE of state (b_1, b_2)
    to find an expression for $\widetilde{\theta}(b_1, b_2)$ as a function of κ
  - **(e)** use the GBE of state $(b_1 - 1, b_2)$
    to find an expression for $\widetilde{\theta}(b_1 - 1, b_2 - 1)$ as a function of κ
  - **(f)** use the GBE of state $(b_1, b_2 - 1)$ to solve for κ
  - **(g)** substitute the value of κ into the equations in the above steps **(1)(a)** and **(1)(c)-(e)**

  (2) if $b_2 > k_2 \geq 2$ (i.e. $b_1, b_2 \geq 3$),

  - **(a)** set $\widetilde{\theta}(0, k_2) = \kappa$
  - **(b)** for $k_1 = 0, \ldots, b_1 - (b_2 - k_2) - 1$
    - **(i)** if $k_1 < b_1 - (b_2 - k_2) - 1$,
      use the GBE of state (k_1, k_2)
      to find an expression for $\widetilde{\theta}(k_1 + 1, k_2)$ as a function of κ
    - **(ii)** if $k_1 = b_1 - (b_2 - k_2) - 1$,
      use the GBE of state (k_1, k_2)
      to find an expression for $\widetilde{\theta}(k_1, k_2 - 1)$ as a function of κ
  - **(c)** for $\ell = k_2, \ldots, 1$
    use the GBE of state $(b_1 - (b_2 - k_2), \ell)$
    to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2), \ell - 1)$ as a function of κ
  - **(d)** use the GBE of state $(b_1 - (b_2 - k_2), 0)$ to solve for κ
  - **(e)** substitute the value of κ into the equations in the above steps **(2)(a)-(c)**

(3) if $b_2 > k_2 = 1$

  (a) set $\widetilde{\theta}(0,1) = \kappa$

  (b) for $k_1 = 0, \dots, b_1 - b_2 + 1$

    (i) if $k_1 < b_1 - b_2$ (if $b_1 - b_2 > 0$),
    use the GBE of state $(k_1, 1)$
    to find an expression for $\widetilde{\theta}(k_1 + 1, 1)$ as a function of κ

    (ii) if $k_1 \in \{b_1 - b_2, b_1 - b_2 + 1\}$,
    use the GBE of state $(k_1, 1)$
    to find an expression for $\widetilde{\theta}(k_1, 0)$ as a function of κ

  (c) for $k_1 = b_1 - b_2, \dots, 1$ (if $b_1 - b_2 > 0$),
  use the GBE of state $(k_1, 0)$
  to find an expression for $\widetilde{\theta}(k_1 - 1, 0)$

  (d) use the GBE of state $(b_1 - (b_2 - 1), 0)$ to solve for κ

  (e) substitute the value of κ into the equations in the above steps **(3)(a)-(c)**

- Normalise all $\widetilde{\theta}(k_1, k_2)$ by setting

$$\widetilde{\theta}(k_1, k_2) \leftarrow \frac{\widetilde{\theta}(k_1, k_2)}{\sum_{k_1=0}^{b_1} \sum_{k_2=0}^{b_2} \widetilde{\theta}(k_1, k_2)}$$

Explanations of the algorithm

Inside the following equations, we use cyan colour for the expressions which are known independent of κ, green colour for the expressions which are known as a function of κ and red colour for the expressions which are unknown. The respective explanations are marked by bullets of the same colour and the expressions are partly highlighted within the explanations. Furthermore, some changes within the explanations are underlined. Furthermore, to support the explanations on the following pages some steps of the algorithm are visualised on the right sides.

We illustrate the algorithm by a simple example with two locations in Appendix C.1.

The algorithm starts by assigning some arbitrary initial value to state $(b_1, 0)$.
The algorithm distinguishes between three cases:
(1) $k_2 = b_2$, **(2)** $b_2 > k_2 \geq 2$ and **(3)** $b_2 > k_2 = 1$.

(1) The case, where $\underline{k_2 = b_2}$, starts with step **(1)(a)** by temporarily setting a variable κ to the prenormalised probability of $(0, b_2)$ and then in step **(1)(b)-(1)(e)**, prenormalised probabilities can be calculated recursively independent of κ resp. as a function of κ by using GBEs. In step **(1)(f)**, the GBE of state $(b_1, b_2 - 1)$ can be used to solve for the variable κ. Hence, in step **(1)(g)**, the value of κ can be substituted into the prenormalised probabilities, which were calculated in the previous steps **(1)(a)** and **(1)(c)-(e)** as a function of κ.

(1)(b) The recursion starts with $\underline{\ell = 0}$.
The GBE of state $(b_1, 0)$ is

$$\underbrace{\widetilde{\theta}(b_1,0)}_{=1} \cdot \Big(\lambda_1 \cdot \underbrace{1_{\{k_1>0\}}}_{=1} + \underbrace{\sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}(b_1,1) \cdot \lambda_2 \cdot \underbrace{1_{\{1<b_2\}}}_{=1}.$$

- $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{2\}$, but a flux into the state $(b_1, 0)$ through a replenishment at location 2 is not possible.
- $\widetilde{\theta}(b_1,0)$ was set to 1 at the beginning of the algorithm.
- Hence, this GBE can be used to solve for $\widetilde{\theta}(b_1,1)$ independent of κ.

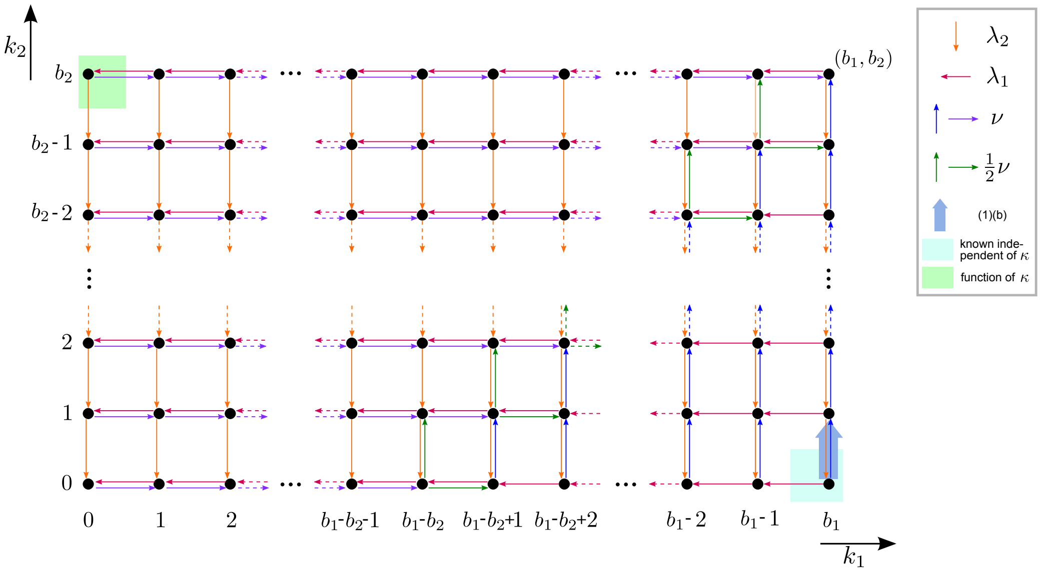
In the next steps, we have $\underline{\ell = 1, \ldots, b_2 - 2}$.
The GBE of state (b_1, ℓ) is

$$\widetilde{\theta}(b_1,\ell) \cdot \Big(\lambda_1 \cdot \underbrace{1_{\{b_1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{\ell>0\}}}_{=1} + \underbrace{\sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

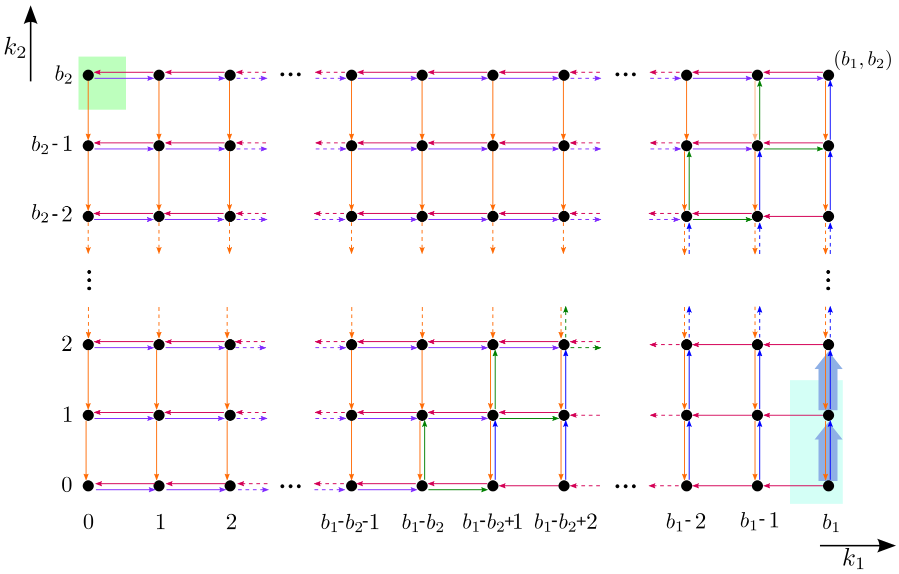
$$= \widetilde{\theta}(b_1,\ell+1) \cdot \lambda_2 \cdot \underbrace{1_{\{\ell<b_2\}}}_{=1}$$

$$+\widetilde{\theta}(b_1,\ell-1) \cdot \nu \cdot \underbrace{p_2(\mathbf{k}-\mathbf{e}_2)}_{=1} \cdot \underbrace{1_{\{\ell>0\}}}_{=1}.$$

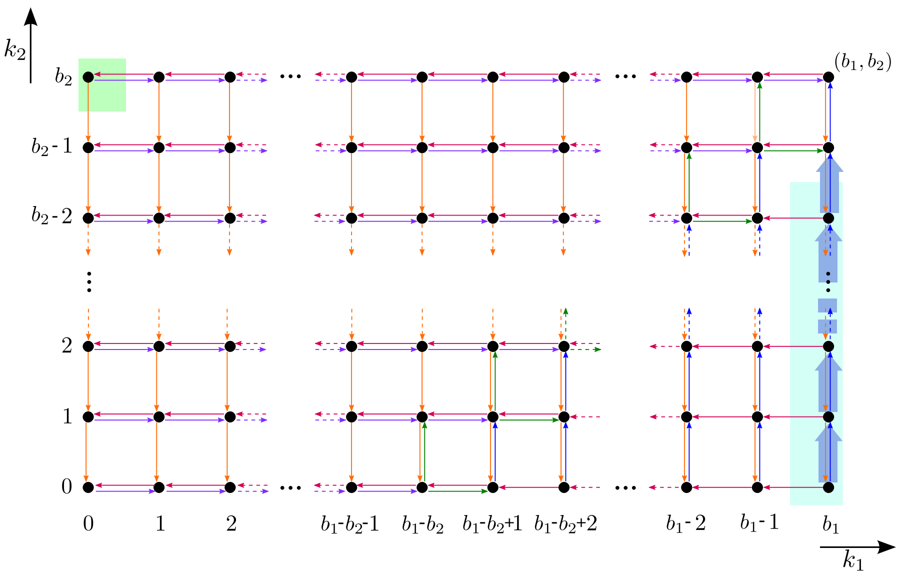
- $p_2(\mathbf{k} - \mathbf{e}_2) = 1$ holds for the following reason:
  Due to $0 < \ell < b_2 - 1$ and $k_1 = b_1$, we have $b_1 - k_1 = 0 < b_2 - (\ell - 1)$, which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{2\}$.
- $\widetilde{\theta}(b_1,\ell)$ for $0 < \ell < b_2 - 1$ is known independent of κ from the previous steps because in step **(1)(b)** the loop of ℓ goes from 0 to $b_2 - 2$.
  More precisely, in the previous step, the GBE of state $(b_1, \underline{\ell - 1})$ was used to find an expression for $\widetilde{\theta}(b_1,\underline{\ell})$.
- $\widetilde{\theta}(b_1,\ell-1)$ is known independent of κ from the previous steps. More precisely,
  - if $\ell = 1$, because $\widetilde{\theta}(b_1,0)$ was set to 1 at the beginning of the algorithm,
  - if $\ell > 1$, because in step **(1)(b)** the loop of ℓ goes from 0 to $b_2 - 2$.
    The GBE of state $(b_1, \underline{\ell - 2})$ was used to find an expression for $\widetilde{\theta}(b_1, \underline{\ell - 1})$.
- Hence, this GBE can be used to solve for $\widetilde{\theta}(b_1,\ell+1)$ independent of κ.



(a) Step **(1)(b)** with $\ell = 0$



(b) Step **(1)(b)** with $\ell = 1$



(c) Step **(1)(b)** with $\ell = b_2 - 2$

Figure C.1.3.: Visualisation of step **(1)(b)** of the algorithm

(1)(c) The recursion starts with $\underline{k_1 = 0}$.
The GBE of state $(0, b_2)$ is

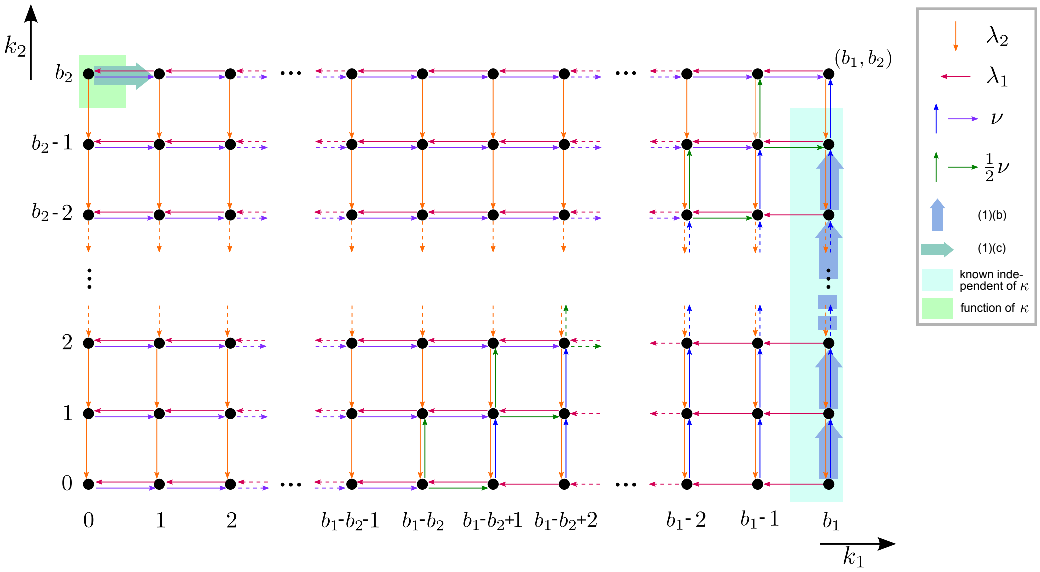
$$\underbrace{\widetilde{\theta}\,(0, b_2)}_{=\kappa} \cdot \left(\lambda_2 \cdot \underbrace{1_{\{b_2>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\right)$$
$$= \widetilde{\theta}\,(1, b_2) \cdot \lambda_1 \cdot \underbrace{1_{\{0<b_1\}}}_{=1}.$$

- $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$, but a flux into the state $(0, b_2)$ through a replenishment is not possible because $1_{\{k_1>0\}} = 0$.
- $\widetilde{\theta}\,(0, b_2)$ was set to κ in step **(1)(a)**.
- Hence, this GBE can be used to solve for $\widetilde{\theta}\,(1, b_2)$ as a function of κ.

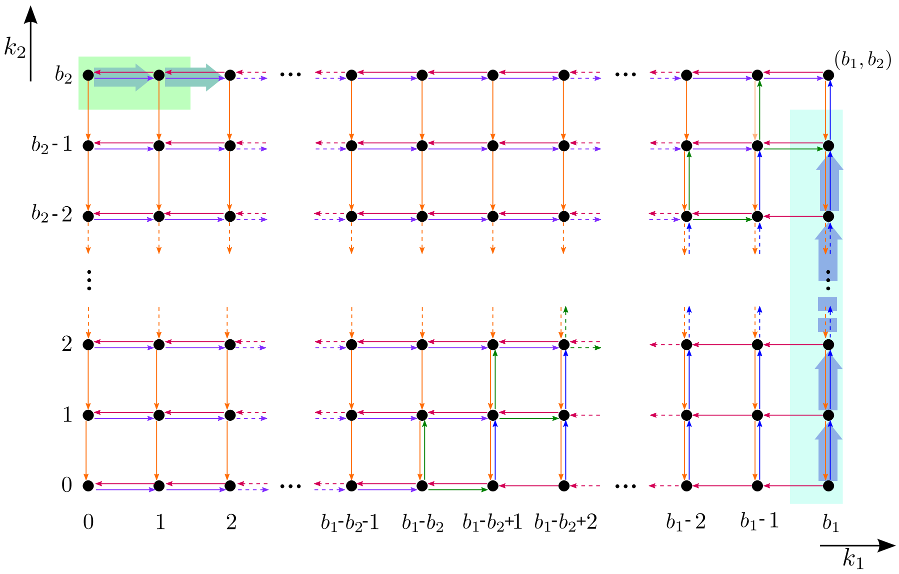
In the next steps, we have $\underline{k_1 = 1, \ldots, b_1 - 2}$.
The GBE of state (k_1, b_2) is

$$\widetilde{\theta}\,(k_1, b_2) \cdot \left(\lambda_1 \cdot \underbrace{1_{\{k_1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{b_2>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\right)$$
$$= \widetilde{\theta}\,(k_1 + 1, b_2) \cdot \lambda_1 \cdot \underbrace{1_{\{k_1<b_1\}}}_{=1}$$
$$+ \widetilde{\theta}\,(k_1 - 1, b_2) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{k_1>0\}}}_{=1}.$$

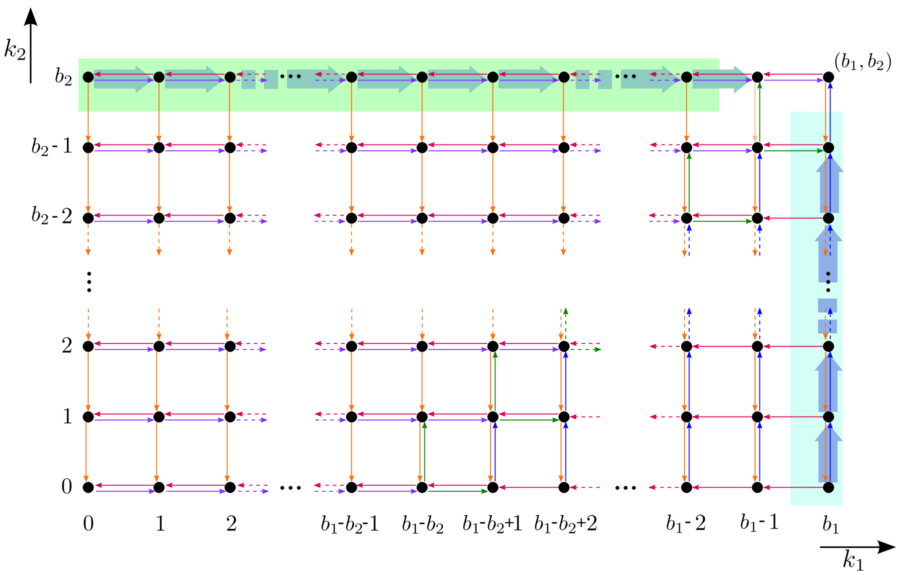
- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $0 < k_1 < b_1$ and $k_2 = b_2$, we have $b_2 - k_2 = 0 < b_1 - (k_1 - 1)$, which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$.
- $\widetilde{\theta}\,(k_1, b_2)$ is known as a function of κ from the previous steps because in step **(1)(c)**, the loop of k_1 goes from 0 to $b_1 - 2$.
  More precisely, because $k_1 > 0$, in step **(1)(c)**, the GBE of state $(\underline{k_1 - 1}, b_2)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, b_2)$.
- $\widetilde{\theta}\,(k_1 - 1, b_2)$ is known for $k_1 > 0$ as a function of κ from the previous steps. More precisely,
  - if $k_1 = 1$, because $\widetilde{\theta}(0, b_2)$ was set to κ in step **(1)(a),**
  - if $k_1 > 1$, because in step **(1)(c)**, the loop of k_1 goes from 0 to $b_1 - 2$.
    The GBE of state $(\underline{k_1 - 2}, b_2)$ was used
    to find an expression for $\widetilde{\theta}(\underline{k_1 - 1}, b_2)$.
- Therefore, this GBE can be used to solve for $\widetilde{\theta}\,(k_1 + 1, b_2)$ as a function of κ.



(a) Step **(1)(c)** with $k_1 = 0$



(b) Step **(1)(c)** with $k_1 = 1$



(c) Step **(1)(c)** with $k_1 = b_1 - 2$

Figure C.1.4.: Visualisation of step **(1)(c)** of the algorithm

(1)(d) The GBE of state (b_1, b_2) is

$$\begin{aligned}
&\widetilde{\theta}(b_1, b_2) \cdot \Big(\lambda_1 \cdot \underbrace{1_{\{b_1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{b_2>0\}}}_{=1}\Big) \\
= &\widetilde{\theta}(b_1 - 1, b_2) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{b_1>0\}}}_{=1} \\
&+ \widetilde{\theta}(b_1, b_2 - 1) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1} \cdot \underbrace{1_{\{b_2>0\}}}_{=1}.
\end{aligned}$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 = b_1$ and $k_2 = b_2$, we have $b_1 - (k_1 - 1) = b_1 - (b_1 - 1) = 1 > 0 = b_2 - k_2$, which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{1\}$.

- $p_2(\mathbf{k} - \mathbf{e}_2) = 1$ holds for the following reason:
  Due to $k_1 = b_1$ and $k_2 = b_2$, we have $b_2 - (k_2 - 1) = b_2 - (b_2 - 1) = 1 > 0 = b_1 - k_1$, which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{2\}$.

- $\widetilde{\theta}(b_1 - 1, b_2)$ is known as a function of κ from the previous steps because in step **(1)(c)**, the loop of k_1 goes from 0 to $b_1 - 2$. More precisely, because $b_1 > 1$, in step **(1)(c)**, the GBE of state $(\underline{b_1 - 2}, b_2)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - 1}, b_2)$.

- $\widetilde{\theta}(b_1, b_2 - 1)$ is known for $b_2 > 1$ independent of κ from the previous steps. More precisely, because in step **(1)(b),** the loop of ℓ goes from 0 to $b_2 - 2$, the GBE of state $(b_1, \underline{b_2 - 2})$ was used to find an expression for $\widetilde{\theta}(b_1, \underline{b_2 - 1})$.

- Therefore, this GBE can be used to solve for $\widetilde{\theta}(b_1, b_2)$ as a function of κ.

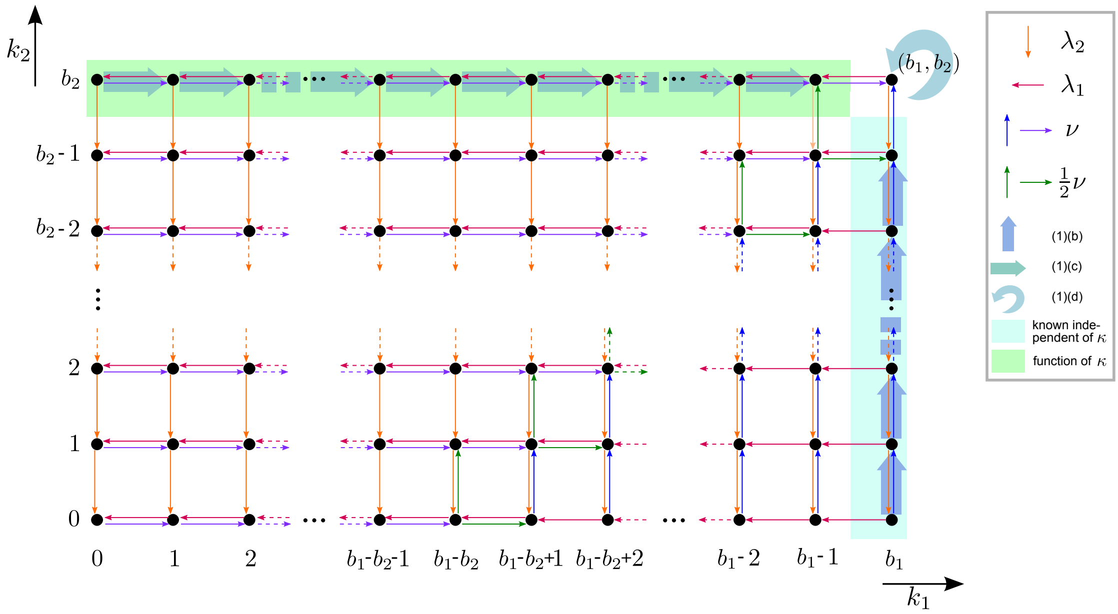


Figure C.1.5.: Visualisation of step **(1)(d)** of the algorithm

(1)(e) The GBE of state $(b_1 - 1, b_2)$ is

$$\widetilde{\theta}(b_1 - 1, b_2) \cdot \Big(\lambda_1 \cdot \underbrace{1_{\{b_1-1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{b_2>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}(b_1, b_2) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1-1<b_1\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - 2, b_2) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{b_1-1>0\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - 1, b_2 - 1) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1/2} \cdot \underbrace{1_{\{b_2>0\}}}_{=1}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - 1$ and $k_2 = b_2$, we have $b_1 - (k_1 - 1) = b_1 - (b_1 - 2) = 2 > 0 = b_2 - k_2$, which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$.

- $p_2(\mathbf{k} - \mathbf{e}_2) = 1/2$ holds for the following reason:
  Due to $k_1 = b_1 - 1$ and $k_2 = b_2$, we have $b_2 - (k_2 - 1) = b_2 - (b_2 - 1) = 1 = b_1 - (b_1 - 1) = b_1 - k_1$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1, 2\}$.

- $\widetilde{\theta}(b_1 - 1, b_2)$ is known for $b_1 > 1$ as a function of κ from the previous steps because in step **(1)(c)**, the loop of k_1 goes from 0 to $b_1 - 2$. More precisely, in step **(1)(c)**, the GBE of state $(b_1 - 2, b_2)$ was used to find an expression for $\widetilde{\theta}(b_1 - 1, b_2)$.

- $\widetilde{\theta}(b_1, b_2)$ is known as a function of κ from the previous step **(1)(d)**.

- $\widetilde{\theta}(b_1 - 2, b_2)$ is known as a function of κ from the previous steps. More precisely,
  - if $b_1 = 2$, because in step **(1)(a)** $\widetilde{\theta}(0, b_2)$ was set to κ,
  - if $b_1 > 2$, because in step **(1)(c)**, the loop of k_1 goes from 0 to $b_1 - 2$.
    The GBE of state $(b_1 - 3, b_2)$ was used
    to find an expression for $\widetilde{\theta}(b_1 - 2, b_2)$.

- Therefore, this GBE can be used to solve for $\widetilde{\theta}(b_1 - 1, b_2 - 1)$ as a function of κ.

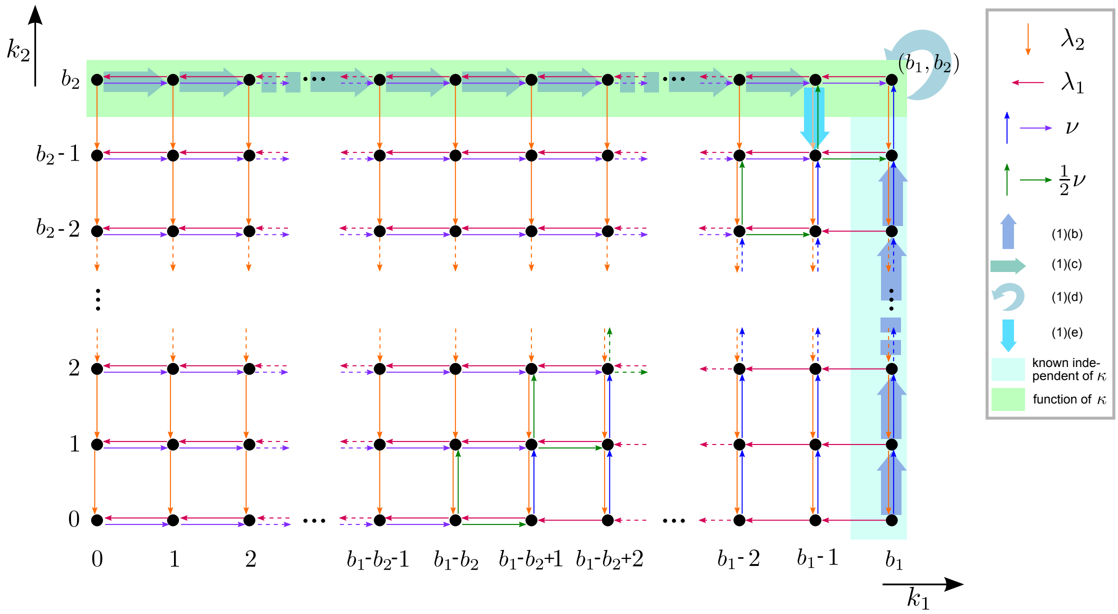


Figure C.1.6.: Visualisation of step **(1)(e)** of the algorithm

(1)(f) The GBE of state $(b_1, b_2 - 1)$ is

$$\widetilde{\theta}(b_1, b_2 - 1) \cdot \Big(\lambda_1 \cdot \underbrace{1_{\{b_1>0\}}}_{=1} + \underbrace{\lambda_2 \cdot 1_{\{b_2-1>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}(b_1, b_2) \cdot \lambda_2 \cdot \underbrace{1_{\{b_2-1<b_2\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - 1, b_2 - 1) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1/2} \cdot \underbrace{1_{\{b_1>0\}}}_{=1}$$

$$+\widetilde{\theta}(b_1, b_2 - 2) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1} \cdot \underbrace{1_{\{b_2-1>0\}}}_{=1} .$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1/2$ holds for the following reason:
  Due to $k_1 = b_1$ and $k_2 = b_2 - 1$, we have
  $b_1 - (k_1 - 1) = b_1 - (b_1 - 1) = 1 = b_2 - (b_2 - 1) = b_2 - k_2$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1, 2\}$.

- $p_2(\mathbf{k} - \mathbf{e}_2) = 1$ holds for the following reason:
  Due to $k_1 = b_1$ and $k_2 = b_2 - 1$, we have $b_2 - (k_2 - 1) = b_2 - (b_2 - 2) = 2 > 0 = b_1 - k_1$, which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{2\}$.

- $\widetilde{\theta}(b_1, b_2 - 1)$ is known for $b_2 > 1$ independent of κ from the previous steps. More precisely, because in step **(1)(b)** the loop of ℓ goes from 0 to $b_2 - 2$, the GBE of state $(b_1, \underline{b_2 - 2})$ was used to find an expression for $\widetilde{\theta}(b_1, \underline{b_2 - 1})$.

- $\widetilde{\theta}(b_1, b_2)$ is known as a function of κ from the previous step **(1)(d)**.

- $\widetilde{\theta}(b_1 - 1, b_2 - 1)$ is known as a function of κ from the previous step **(1)(e)**.

- $\widetilde{\theta}(b_1, b_2 - 2)$ is known independent of κ from the previous steps. More precisely,
  - if $b_2 = 2$, because $\widetilde{\theta}(b_1, 0)$ was set to 1 at the beginning of the algorithm.
  - if $b_2 > 2$, because in step **(1)(b)**, the loop of ℓ goes from 0 to $b_2 - 2$. The GBE of state $(b_1, \underline{b_2 - 3})$ was used to find an expression for $\widetilde{\theta}(b_1, \underline{b_2 - 2})$.

Hence, everything is known and at least one expression is independent of κ.
Thus, we can use this GBE to solve for κ.

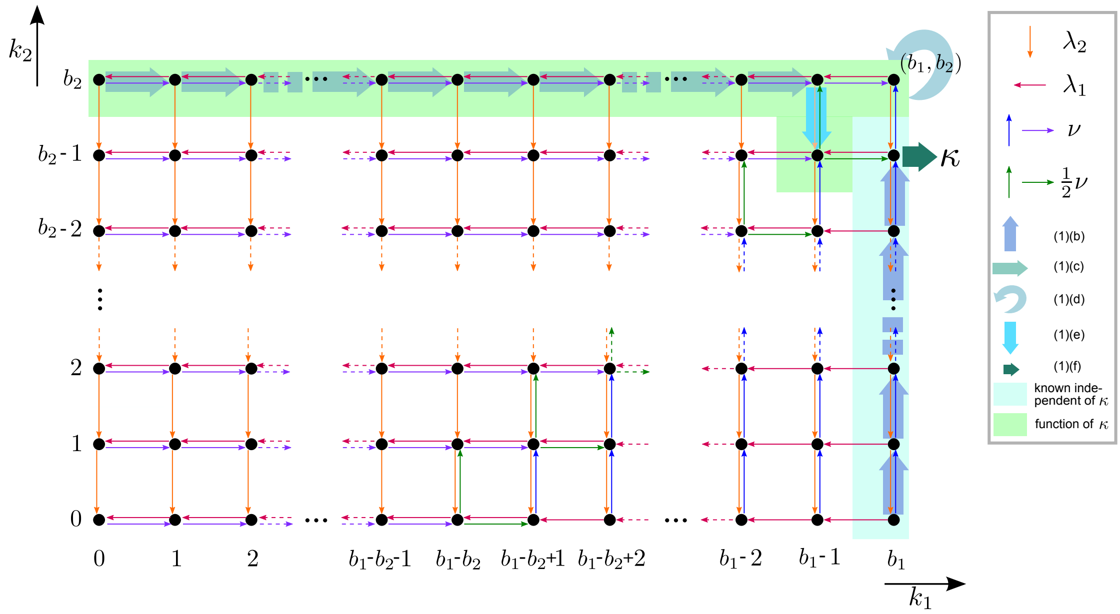


Figure C.1.7.: Visualisation of step **(1)(f)** of the algorithm

(2) The case, where $\underline{b_2 > k_2 \geq 2}$ (this implies that $b_1 \geq 3$), starts with step **(2)(a)** by temporarily setting a variable κ to the prenormalised probability of $(0, k_2)$, and then in step **(2)(b)** and **(2)(c)**, prenormalised probabilities can be calculated recursively as a function of κ by using GBEs. In step **(2)(d)**, the GBE of state $(b_1 - (b_2 - k_2), 0)$ can be used to solve for the variable κ. Hence, in step **(2)(e)**, the value of κ can be substituted into the prenormalised probabilities, which were calculated in the previous steps **(2)(a)**-**(c)** as a function of κ.

(2)(b)(i) The recursion starts with $\underline{0 = k_1 < b_1 - (b_2 - k_2) - 1}$.
The GBE of state $(0, k_2)$ is

$$\underbrace{\widetilde{\theta}\,(0, k_2)}_{=\kappa} \cdot \Big(\lambda_2 \cdot \underbrace{1_{\{k_2>0\}}}_{=1} + \underbrace{\sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$
$$= \widetilde{\theta}\,(0, k_2+1) \cdot \lambda_2 \cdot \underbrace{1_{\{k_2<b_2\}}}_{=1}$$
$$+ \widetilde{\theta}\,(1, k_2) \cdot \lambda_1 \cdot \underbrace{1_{\{0<b_1\}}}_{=1} .$$

- $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{1\}$, but a flux into the state $(0, k_2)$ through a replenishment at location 1 is not possible because $1_{\{k_1>0\}} = 0$.
- $\widetilde{\theta}\,(0, k_2)$ was set to κ in step **(2)(a)**.
- $\widetilde{\theta}(0, k_2 + 1)$ is known for $k_2 < b_2$ independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $k_2 = b_2 - 1$, because $\widetilde{\theta}(0, b_2)$ was calculated in step **(1)**.
    More precisely, in step **(1)(a)**, $\widetilde{\theta}(0, b_2)$ was set to κ and in step **(1)(g)**, the value of κ was substituted.
  - if $k_2 < b_2 - 1$, because $\widetilde{\theta}(0, k_2 + 1)$ was calculated in step **(2)**.
    More precisely, in step **(2)(a)**, $\widetilde{\theta}(0, k_2 + 1)$ was set to κ and in step **(2)(e)**, the value of κ was substituted.
- Hence, this GBE can be used to solve for $\widetilde{\theta}\,(1, k_2)$ as a function of κ.

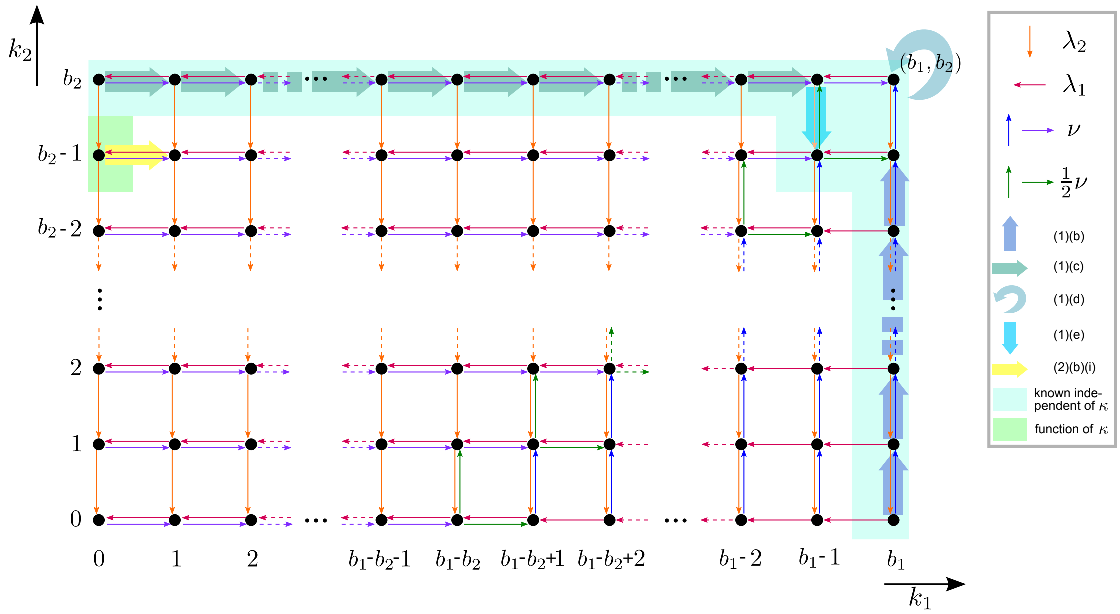


Figure C.1.8.: Visualisation of step **(2)(b)(i)** with $k_2 = b_2 - 1$ and $k_1 = 0$ of the algorithm

In the next steps, we have $\underline{0 < k_1 < b_1 - (b_2 - k_2) - 1}$.
The GBE of state (k_1, k_2) is

$$\widetilde{\theta}\,(k_1, k_2) \cdot \Big(\lambda_1 \cdot \underbrace{1_{\{k_1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{k_2>0\}}}_{=1} + \underbrace{\sum_{i\in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}\,(k_1, k_2 + 1) \cdot \lambda_2 \cdot \underbrace{1_{\{k_2<b_2\}}}_{=1}$$

$$+\widetilde{\theta}\,(k_1 + 1, k_2) \cdot \lambda_1 \cdot \underbrace{1_{\{k_1<b_1\}}}_{=1}$$

$$+\widetilde{\theta}\,(k_1 - 1, k_2) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{k_1>0\}}}_{=1} .$$

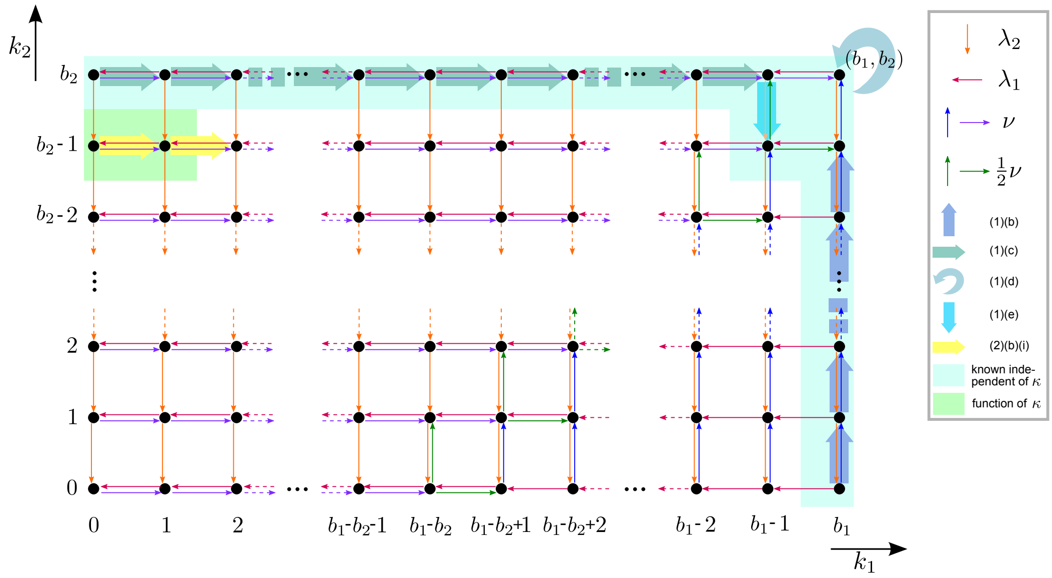
- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $0 < b_2 - k_2 < b_1 - k_1 - 1$, we have $b_2 - k_2 < b_1 - (k_1 - 1)$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$.

- $\widetilde{\theta}\,(k_1, k_2)$ is known as a function of κ from the previous steps because in step **(2)(b)**, the loop of k_1 goes from 0 to $b_1 - (b_2 - k_2) - 1$.
  More precisely, because $k_1 > 0$, in step **(2)(b)(i)**, the GBE of state $(\underline{k_1 - 1}, k_2)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, k_2)$.

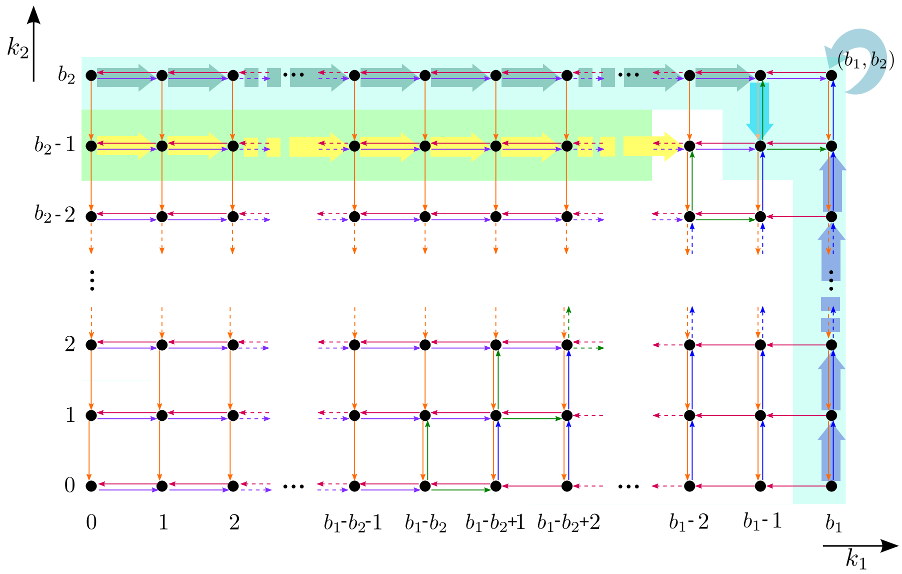
- $\widetilde{\theta}(k_1, k_2 + 1)$ is known for $k_2 < b_2$ independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $k_2 = b_2 - 1$, because in step **(1)(c)**, the GBE of state $(\underline{k_1 - 1}, b_2)$ was used to find an expression for $\widetilde{\theta}(k_1, \underline{b_2})$.
  - if $k_2 < b_2 - 1$, because in step **(2)(b)(i)**, the GBE of state $(\underline{k_1 - 1}, k_2 + 1)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, k_2 + 1)$.

- $\widetilde{\theta}\,(k_1 - 1, k_2)$ is known for $k_1 > 0$ as a function of κ from the previous steps.
  More precisely,
  - if $k_1 = 1$, because $\widetilde{\theta}\,(k_1 - 1, k_2) = \widetilde{\theta}(0, k_2)$ was set to κ in step **(2)(a)**,
  - if $k_1 > 1$, because in step **(2)(b)**, the loop of k_1 goes from 0 to $b_1 - (b_2 - k_2) - 1$.
    In step **(2)(b)(i)**, the GBE of state $(\underline{k_1 - 2}, k_2)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1 - 1}, k_2)$.

- Therefore, this GBE can be used to solve for $\widetilde{\theta}\,(k_1 + 1, k_2)$ as a function of κ.



(a) Step **(2)(b)(i)** with $k_2 = b_2 - 2$ and $k_1 = 2$



(b) Step **(2)(b)(i)** with $k_2 = b_2 - 2$ and $k_1 = b_1 - 3$

Figure C.1.9.: Visualisation of step **(2)(b)(i)** of the algorithm

(2)(b)(ii) If $k_1 = b_1 - (b_2 - k_2) - 1$, then we have $b_1 - k_1 = b_2 - k_2 + 1$.
The GBE of state $(k_1, k_2) = (b_1 - (b_2 - k_2) - 1, k_2)$ is

$$\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 1, k_2)\Big(\lambda_1 \cdot \underbrace{1_{\{b_1-(b_2-k_2)-1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{k_2>0\}}}_{=1} + \underbrace{\sum_{i\in\bar{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}\,(b_1 - (b_2 - k_2) - 1, k_2 + 1) \cdot \lambda_2 \cdot \underbrace{1_{\{k_2<b_2\}}}_{=1}$$

$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2), k_2) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1-(b_2-k_2)-1<b_1\}}}_{=1}$$

$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 2, k_2) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot 1_{\{b_1-(b_2-k_2)-1>0\}}$$

$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 1, k_2 - 1) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1/2} \cdot 1_{\{k_2>0\}}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - (b_2 - k_2) - 1$, we have
  $b_1 - (k_1 - 1) = b_1 - (b_1 - (b_2 - k_2) - 2) = (b_2 - k_2) + 2 > b_2 - k_2$,
  which leads to $\arg\max_{j\in\bar{J}}(b_j - k_j) = \{1\}$.
- $p_2(\mathbf{k} - \mathbf{e}_2) = 1/2$ holds for the following reason:
  Due to $k_1 = b_1 - (b_2 - k_2) - 1$,
  we have $b_1 - k_1 = b_1 - (b_1 - (b_2 - k_2) - 1) = b_2 - k_2 + 1 = b_2 - (k_2 - 1)$,
  which leads to $\arg\max_{j\in\bar{J}}(b_j - k_j) = \{1, 2\}$.
- $\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 1, k_2)$ is known as a function of κ from the previous steps because in step **(2)(b)** the loop of k_1 goes from 0 to $b_1 - (b_2 - k_2) - 1$. More precisely, in step **(2)(b)(i)**, the GBE of state $(b_1 - (b_2 - k_2) - 2, k_2)$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2) - 1, k_2)$.

- $\widetilde{\theta}(b_1 - (b_2 - k_2) - 1, k_2 + 1)$ is known for $k_2 < b_2$ independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $k_2 = b_2 - 1$, because in step **(1)(c)**, the GBE of state $(b_1 - 3, b_2)$ was used to find an expression for
    $\widetilde{\theta}(b_1 - (b_2 - b_2) - 1, b_2) = \widetilde{\theta}(b_1 - 2, b_2)$.
  - if $k_1 < b_2 - 1$, because in step **(2)(b)(i)**, the GBE of state
    $(b_1 - (b_2 - (k_2 + 1)) - 3, k_2 + 1) = (b_1 - (b_2 - k_2) - 2, k_2 + 1)$
    was used to find an expression for
    $\widetilde{\theta}(b_1 - (b_2 - (k_2 + 1)) - 2, k_2+1) = \widetilde{\theta}(b_1 - (b_2 - k_2) - 1, k_2 + 1)$.

- $\widetilde{\theta}\,(b_1 - (b_2 - k_2), k_2)$ is known for $b_1 - (b_2 - k_2) - 1 < b_1$ independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $k_2 = b_2 - 1$, because $b_1 - (b_2 - k_2) = b_1 - (b_2 - (b_2 - 1)) = b_2 - 1$,
    in step **(1)(e)**, the GBE of state $(b_1 - 1, b_2)$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2), k_2) = \widetilde{\theta}(b_1 - 1, b_2 - 1)$,

- if $k_2 < b_2 - 1$, because $b_1 - (b_2 - k_2) = b_1 - (b_2 - (k_2+1)) - 1$, in step **(2)(b)(ii)**, the GBE of state $(b_1 - (b_2 - (k_2+1)) - 1, \underline{k_2+1})$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - (k_2+1)) - 1, \underline{k_2}) = \widetilde{\theta}(b_1 - (b_2 - k_2), \underline{k_2})$.

- $\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 2, k_2)$ is known for $b_1 - (b_2 - k_2) - 1 > 0$ as a function of κ from the previous steps. More precisely,
  - if $b_1 - (b_2 - k_2) - 2 = 0$, because in step **(2)(a)**, $\widetilde{\theta}(0, k_2)$ was set to κ,
  - if $b_1 - (b_2 - k_2) - 2 > 0$, because in step **(2)(b)** the loop of k_1 goes from 0 to $b_1 - (b_2 - k_2) - 1$. In step **(2)(b)(i)**, the GBE of state $\underline{(b_1 - (b_2 - k_2) - 3}, k_2)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - (b_2 - k_2) - 2}, k_2)$.

- Consequently, this GBE can be used to solve for $\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 1, k_2 - 1)$ as a function of κ.

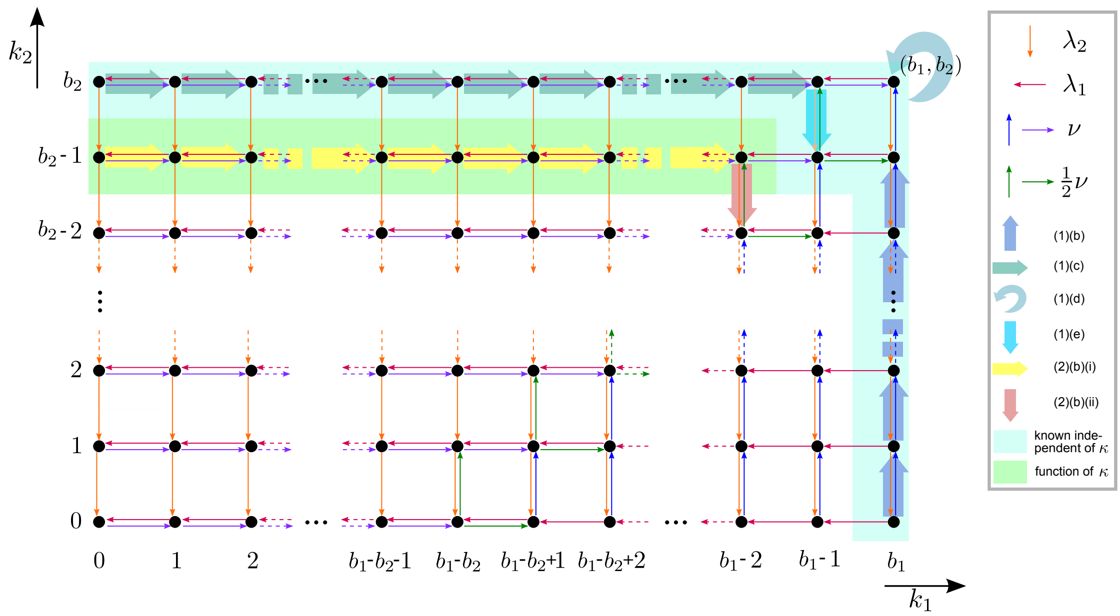


Figure C.1.10.: Visualisation of step **(2)(b)(ii)** with $k_2 = b_2 - 1$ of the algorithm

(2)(c) It starts with $\underline{\ell = k_2}$.
The GBE of the state $(b_1-(b_2-k_2),\ell)=(b_1-(b_2-k_2),k_2)$ is

$$\widetilde{\theta}\,(b_1-(b_2-k_2),k_2)\Big(\lambda_1\cdot \underbrace{1_{\{b_1-(b_2-k_2)>0\}}}_{=1}+\lambda_2\cdot \underbrace{1_{\{k_2>0\}}}_{=1}+\underbrace{\sum_{i\in\overline{J}}\nu\cdot p_i(\mathbf{k})\cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$\begin{aligned}
&=\widetilde{\theta}\,(b_1-(b_2-k_2),k_2+1)\cdot\lambda_2\cdot\underbrace{1_{\{k_2<b_2\}}}_{=1}\\
&\quad+\widetilde{\theta}\,(b_1-(b_2-k_2)+1,k_2)\cdot\lambda_1\cdot\underbrace{1_{\{b_1-(b_2-k_2)<b_1\}}}_{=1}\\
&\quad+\widetilde{\theta}\,(b_1-(b_2-k_2)-1,k_2)\cdot\nu\cdot\underbrace{p_1(\mathbf{k}-\mathbf{e}_1)}_{=1}\cdot\underbrace{1_{\{b_1-(b_2-k_2)>0\}}}_{=1}\\
&\quad+\widetilde{\theta}\,(b_1-(b_2-k_2),k_2-1)\cdot\nu\cdot\underbrace{p_2(\mathbf{k}-\mathbf{e}_2)}_{=1}\cdot\underbrace{1_{\{k_2>0\}}}_{=1}.
\end{aligned}\tag{C.1.1}$$

- $p_1(\mathbf{k}-\mathbf{e}_1)=1$ holds for the following reason:
  Due to $k_1=b_1-(b_2-k_2)$,
  we have $b_1-(k_1-1)=b_1-(b_1-(b_2-k_2)-1)=b_2-k_2+1>b_2-k_2$,
  which leads to $\arg\max\limits_{j\in\overline{J}}(b_j-k_j)=\{1\}$.

- $p_2(\mathbf{k}-\mathbf{e}_2)=1$ holds for the following reason:
  Due to $k_1=b_1-(b_2-k_2)$,
  we have $b_1-k_1=b_1-(b_1-(b_2-k_2))=b_2-k_2<b_2-(k_2-1)$,
  which leads to $\arg\max\limits_{j\in\overline{J}}(b_j-k_j)=\{2\}$.

- $\widetilde{\theta}\,(b_1-(b_2-k_2),k_2)$ is known independent of κ from the previous for-loop because the for-loop of k_2 goes from b_2 to 1. More precisely,
  - if $k_2=b_2-1$, because in step **(1)(e)**, the GBE of state $(b_1-1,\underline{b_2})$ was used to find an expression for $\widetilde{\theta}(b_1-(b_2-k_2),\underline{k_2})=\widetilde{\theta}(b_1-1,\underline{b_2-1})$,
  - if $k_2<b_2-1$, because in step **(2)(b)(ii)**, the GBE of state $(b_1-(b_2-(k_2+1))-1,\underline{k_2+1})$ was used to find an expression for $\widetilde{\theta}(b_1-(b_2-(k_2+1))-1,\underline{k_2})=\widetilde{\theta}(b_1-(b_2-k_2),\underline{k_2})$.

- $\widetilde{\theta}(b_1-(b_2-k_2),k_2+1)$ is known for $k_2<b_2$ independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $k_2=b_2-1$, because in step **(1)(c)**, the GBE of state $(\underline{b_1-(b_2-k_2)}-1,b_2)=(\underline{b_1-2},b_2)$ was used to find an expression for $\widetilde{\theta}(b_1-(b_2-k_2),b_2)=\widetilde{\theta}(\underline{b_1-1},b_2)$.
  - if $k_1<b_2-1$, because in step **(2)(b)(i)**, the GBE of state $(b_1-(b_2-(k_2+1))-2,k_2+1)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1-(b_2-(k_2+1))-1},k_2+1)=\widetilde{\theta}(\underline{b_1-(b_2-k_2)},k_2+1)$.

- $\widetilde{\theta}\,(b_1-(b_2-k_2)+1,k_2)$ are known for $b_1-(b_2-k_2)<b_1$ independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1.
  More precisely,
  - if $k_2=b_2-1$, because $b_1-(b_2-k_2)+1=b_1-(b_2-(b_2-1)+1)=b_1$, in step **(1)(b)**, the GBE of state $(b_1,\underline{b_2-2})$ was used to find an expression for $\widetilde{\theta}(b_1,\underline{b_2-1})$,
  - if $k_2<b_2-1$, because in step **(2)(c)**, the GBE of state $(b_1-(b_2-(k_2+1))\,,\underline{k_2+1})$ was used to find an expression for $\widetilde{\theta}(b_1-(b_2-(k_2+1))\,,\underline{k_2})=\widetilde{\theta}(b_1-(b_2-k_2)+1,\underline{k_2})$.
- $\widetilde{\theta}\,(b_1-(b_2-k_2)-1,k_2)$ is known for $b_1-(b_2-k_2)>0$ as a function of κ from the previous steps because in step **(2)(b)** the loop of k_1 goes from 0 to $b_1-(b_2-k_2)-1$. More precisely, in step **(2)(b)(i)**, the GBE of state $(\underline{b_1-(b_2-k_2)-2},k_2)$ was used to find an expression for $\widetilde{\theta}\,(b_1-(b_2-k_2)-1,k_2)$.
- Consequently, this GBE can be used to solve for $\widetilde{\theta}\,(b_1-(b_2-k_2),k_2-1)$ as a function of κ.

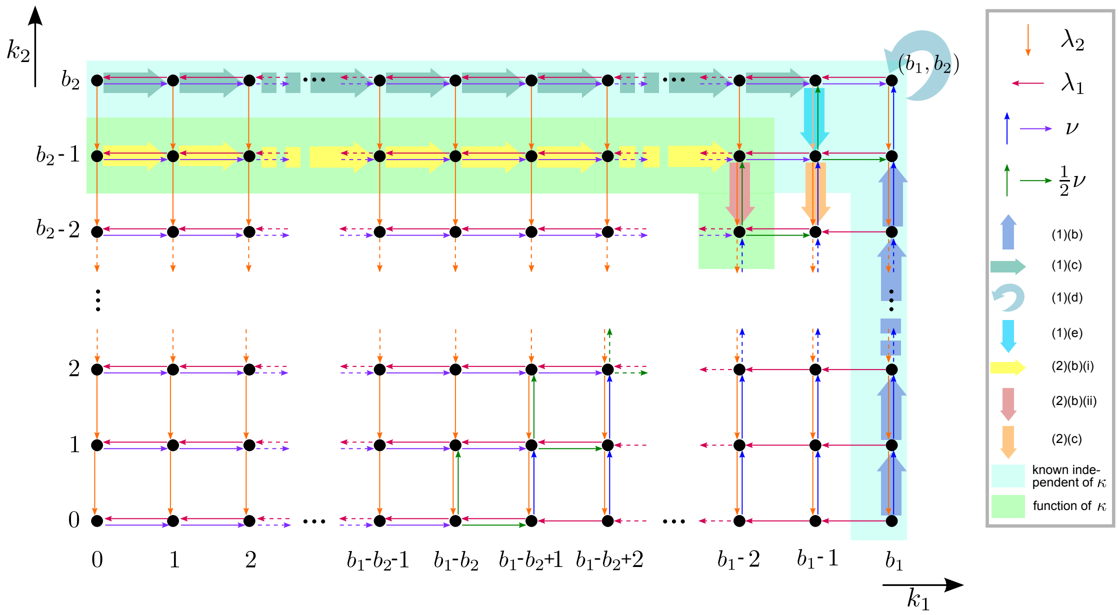


Figure C.1.11.: Visualisation of step **(2)(c)** with $k_2=b_2-1$ and $\ell=k_2$ of the algorithm

In the next steps, where $\underline{k_2 > \ell > 0}$.
The GBE of state $(b_1 - (b_2 - k_2), \ell)$ is

$$\widetilde{\theta}\,(b_1 - (b_2 - k_2), \ell)\Big(\lambda_1 \cdot \underbrace{1_{\{b_1-(b_2-k_2)>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{\ell>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}\,(b_1 - (b_2 - k_2), \ell + 1) \cdot \lambda_2 \cdot \underbrace{1_{\{\ell<b_2\}}}_{=1}$$

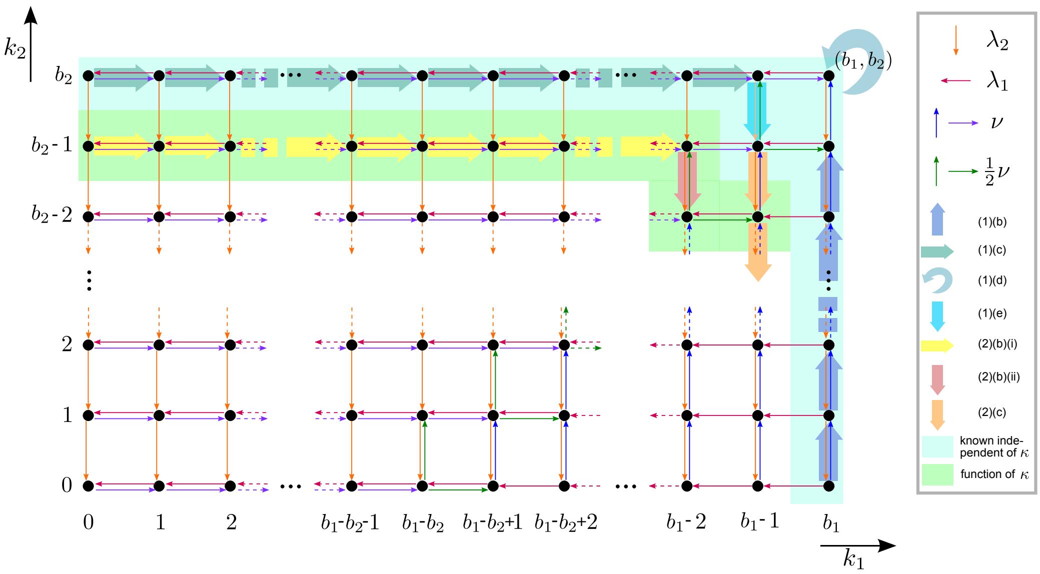
$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2) + 1, \ell) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1-(b_2-k_2)<b_1\}}}_{=1}$$

$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2) - 1, \ell) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1/2} \cdot \underbrace{1_{\{b_2-(k_2-1)=b_2-\ell\}}}_{=1_{\{\ell=k_2-1\}}}$$

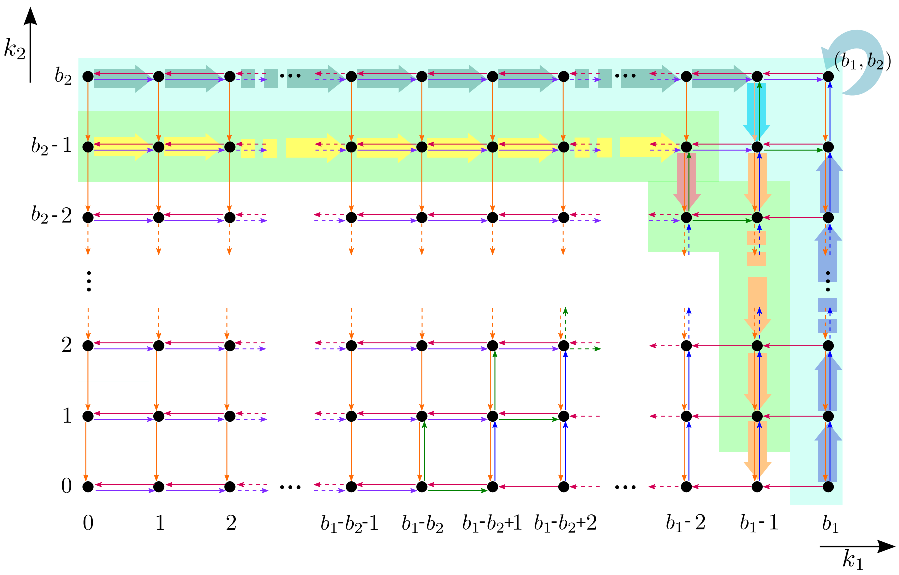
$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2), \ell - 1) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1} \cdot \underbrace{1_{\{\ell>0\}}}_{=1}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1/2$ holds for the following reason:
  Due to $k_1 = b_1 - (b_2 - k_2)$ and if $b_2 - k_2 + 1 = b_2 - \ell$,
  we have $b_1 - (k_1 - 1) = b_1 - (b_1 - (b_2 - k_2) - 1) = b_2 - k_2 + 1 = b_2 - \ell$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1, 2\}$.
- $p_2(\mathbf{k} - \mathbf{e}_2) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - (b_2 - k_2)$ and $k_2 > \ell > 0$,
  we have $b_1 - (b_1 - (b_2 - k_2) = b_2 - k_2 < b_2 - (\ell - 1)$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{2\}$.
- $\widetilde{\theta}\,(b_1 - (b_2 - k_2), \ell)$ is known as a function of κ from the previous steps because in step **(2)(c)** the loop of ℓ goes from k_2 to 1. More precisely, in step **(2)(c)**, the GBE of state $(b_1 - (b_2 - k_2), \underline{\ell + 1})$ was used to find an expression for $\widetilde{\theta}\,(b_1 - (b_2 - k_2), \underline{\ell})$.
- $\widetilde{\theta}\,(b_1 - (b_2 - k_2), \ell + 1)$ is known from the previous steps. More precisely,
  - if $k_2 = b_2 - 1$ and $\ell = k_2 - 1$,
    then $\widetilde{\theta}(b_1 - (b_2 - k_2), \ell + 1) = \widetilde{\theta}(b_1 - 1, b_2 - 1)$ is known independent of κ because in step **(1)(e)**, the GBE of state $(b_1 - 1, \underline{b_2})$ was used to find an expression for $\widetilde{\theta}\,(b_1 - 1, b_2) = \widetilde{\theta}(b_1 - 1, \underline{b_2 - 1})$.
  - if $k_2 < b_2 - 1$ and $\ell = k_2 - 1$,
    then $\widetilde{\theta}(b_1 - (b_2 - k_2), \ell + 1) = \widetilde{\theta}(b_1 - (b_2 - k_2), k_2)$ is known as a function of κ because the for-loop of k_2 goes from b_2 to 1. More precisely, because in step **(2)(b)(ii)**, the GBE of state $(b_1 - (b_2 - (k_2 + 1)) - 1, \underline{k_2 + 1})$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - (k_2 + 1)) - 1, \underline{k_2}) = \widetilde{\theta}(b_1 - (b_2 - k_2), \underline{k_2})$.
  - if $k_2 \leq b_2 - 1$ and $\ell < k_2 - 1$,
    because in step **(2)(c)** the loop of ℓ goes from k_2 to 1. The GBE of state $(b_1 - (b_2 - k_2), \underline{\ell + 2})$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2), \underline{\ell + 1})$.

- $\widetilde{\theta}\,(b_1-(b_2-k_2)+1,\ell)$ is known for $k_2>\ell>0$ independent of κ from a previous for-loop. More precisely,
  - if $k_2=b_2-1$, $\widetilde{\theta}\,(b_1-(b_2-k_2)+1,\ell)=\widetilde{\theta}\,(b_1,\ell)$ is known independent of κ, because in step **(1)(b)** the GBE of state $(b_1,\underline{\ell-1})$ was used to find an expression for $\widetilde{\theta}(b_1,\underline{\ell})$,
  - if $k_2<b_2-1$, is known independent of κ because the loop of k_2 goes from b_2 to 1. Since $b_1-(b_2-k_2)+1=b_1-(b_2-(k_2+1))$, in step **(2)(c)** the GBE of state $(b_1-(b_2-(k_2+1))\,,\underline{\ell+1})$ was used to find an expression for $\widetilde{\theta}(b_1-(b_2-(k_2+1))\,,\underline{\ell})$.
- $\widetilde{\theta}\,(b_1-(b_2-k_2)-1,\ell)$ is known for $\ell=k_2-1$ as a function of κ. More precisely, because in step **(2)(b)(ii)**, the GBE of state $(b_1-(b_2-k_2)-1,\underline{k_2})$ was used to find an expression for $\widetilde{\theta}(b_1-(b_2-k_2)-1,k_2-1)=\widetilde{\theta}(b_1-(b_2-k_2)-1,\underline{\ell})$.
- Consequently, this GBE can be used to solve for $\widetilde{\theta}\,(b_1-(b_2-k_2),\ell-1)$ as a function of κ.



(a) Step **(2)(c)** with $k_2=b_2-1$ and $\ell=k_2-1$



(b) Step **(2)(c)** with $k_2=b_2-1$ and $\ell=1$

Figure C.1.12.: Visualisation of step **(2)(c)** of the algorithm

(2)(d) The GBE of state $(b_1 - (b_2 - k_2), 0)$ is

$$\widetilde{\theta}\,(b_1 - (b_2 - k_2), 0)\Big(\lambda_1 \cdot \underbrace{1_{\{b_1-(b_2-k_2)>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}\,(b_1 - (b_2 - k_2), 1) \cdot \lambda_2 \cdot \underbrace{1_{\{0<b_2\}}}_{=1}$$

$$+\widetilde{\theta}\,(b_1 - (b_2 - k_2) + 1, 0) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1-(b_2-k_2)<b_1\}}}_{=1}$$

- $\widetilde{\theta}\,(b_1 - (b_2 - k_2), 0)$ is known as a function of κ from the previous steps because in step **(2)(c)** the for-loop of ℓ goes from k_2 to 1. More precisely, in step **(2)(c)**, the GBE of state $(b_1 - (b_2 - k_2), \underline{1})$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2), \underline{0})$.
- $\widetilde{\theta}\,(b_1 - (b_2 - k_2), 1)$ is known as a function of κ from the previous steps because in step **(2)(c)** the loop of ℓ goes from k_2 to 1. More precisely, in step **(2)(c)**, the GBE of state $(b_1 - (b_2 - k_2), \underline{2})$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2), \underline{1})$.
- $\widetilde{\theta}\,(b_1 - (b_2 - k_2) + 1, 0)$ is known independent of κ from the previous steps. More precisely,
  - if $k_2 = b_2 - 1$, because $\widetilde{\theta}\,(b_1, 0)$ was set to 1 at the beginning of the algorithm.
  - if $k_2 < b_2 - 1$, because the loop of k_2 goes from b_2 to 1.
    More precisely, because $b_1 - (b_2 - k_2) + 1 = b_1 - (b_2 - (k_2 + 1))$, in step **(2)(c)** the GBE of state $(b_1 - (b_2 - k_2) + 1, \underline{1})$ was used to find an expression for $\widetilde{\theta}(b_1 - (b_2 - k_2) + 1, \underline{0})$.

Hence, everything is known and at least one expression is independent of κ.
Thus, we can use this GBE to solve for κ.

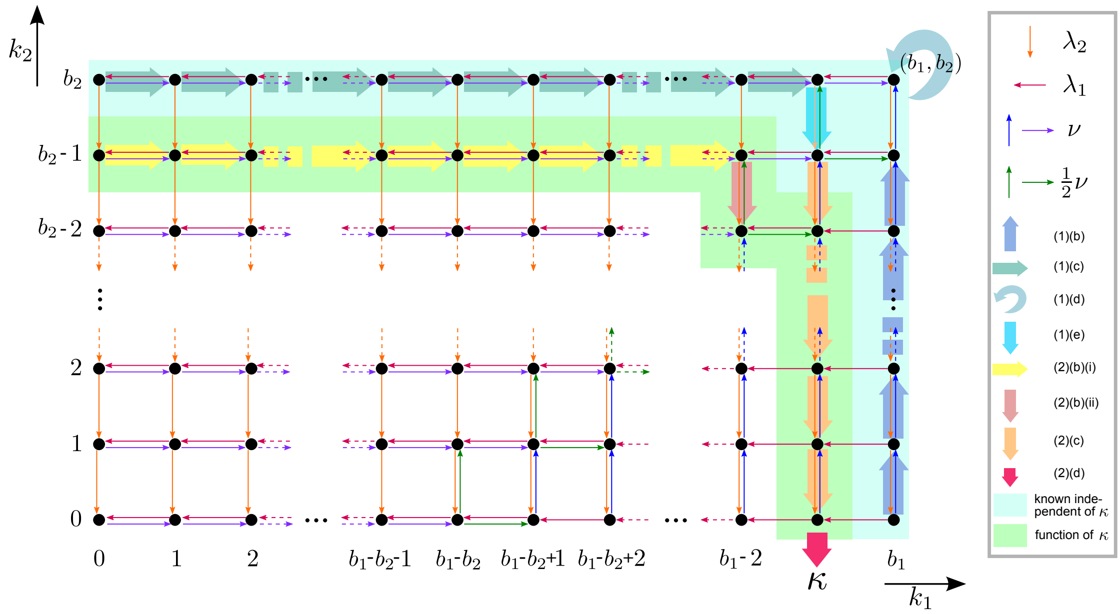


Figure C.1.13.: Visualisation of step **(2)(d)** with $k_2 = b_2 - 1$ of the algorithm

(3) The case, where $\underline{b_2 > k_2 = 1}$ (this implies that $b_1 \geq 3$), starts with step **(3)(a)** by temporarily setting a variable κ to the prenormalised probability of $(0, 1)$, and then in step **(3)(b)** and **(3)(c)**, prenormalised probabilities can be calculated recursively as a function of κ by using GBEs. In step **(3)(d)**, the GBE of state $(b_1 - (b_2 - 1), 0)$ can be used to solve for the variable κ. Hence, in step **(3)(e)**, the value of κ can be substituted into the prenormalised probabilities which were calculated in the previous steps as a function of κ.

(3)(b)(i) The recursion starts with $\underline{0 = k_1 < b_1 - b_2}$.
The GBE of the state $(0, 1)$ is

$$\underbrace{\widetilde{\theta}\,(0,1)}_{=\kappa}\left(\lambda_2 \cdot \underbrace{1_{\{1>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\right)$$

$$= \widetilde{\theta}\,(0,2) \cdot \lambda_2 \cdot 1_{\{1<b_2\}}$$
$$+ \widetilde{\theta}\,(1,1) \cdot \lambda_1 \cdot \underbrace{1_{\{0<b_1\}}}_{=1}.$$

- $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$, but a flux into the state $(0, 1)$ through a replenishment is not possible because $1_{\{k_1>0\}} = 0$.
- $\widetilde{\theta}\,(0,1)$ was set to κ in step **(3)(a)**.
- $\widetilde{\theta}\,(0,2)$ is known for $1 < b_2$ from the previous for-loop independent of κ because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $b_2 = 2$, because $\widetilde{\theta}(0, b_2) = \widetilde{\theta}(0, 2)$ was calculated in step **(1)**.
    More precisely, in step **(1)(a)**, $\widetilde{\theta}(0, 2)$ was set to κ and in step **(1)(g)**, the value of κ was substituted.
  - if $b_2 > 2$, because $\widetilde{\theta}(0, 2)$ was calculated in step **(2)**.
    More precisely, in step **(2)(a)**, $\widetilde{\theta}(0, 2)$ was set to κ and in step **(2)(e)**, the value of κ was substituted.
- Hence, this GBE can be used to solve for $\widetilde{\theta}\,(1,1)$ as a function of κ.

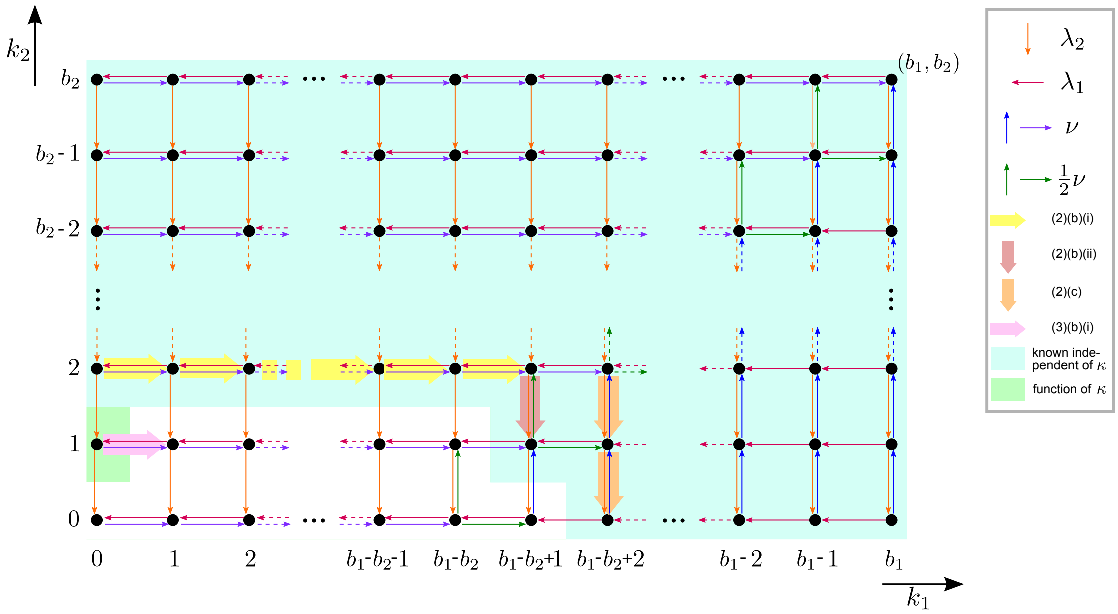
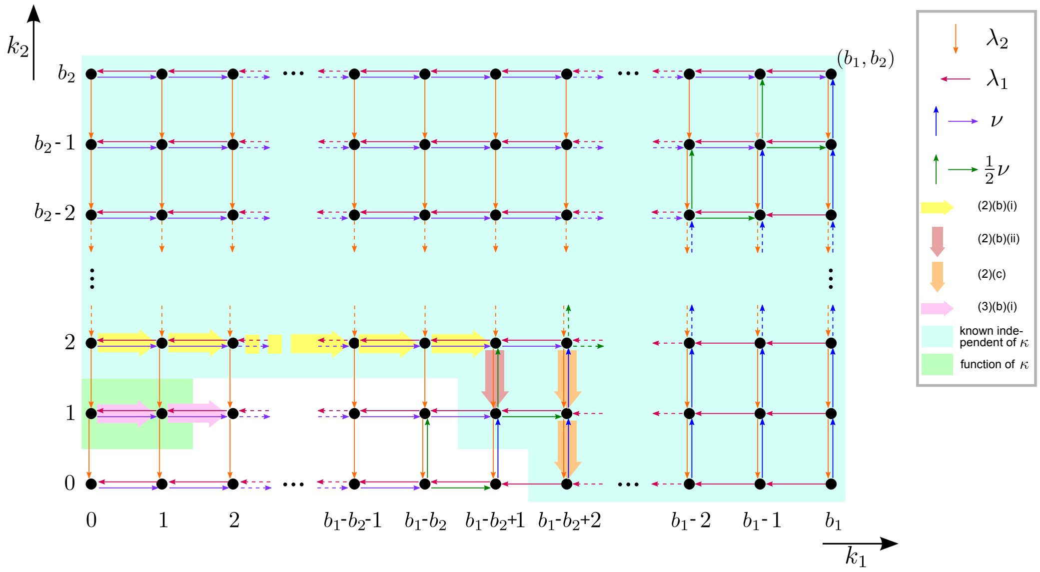


Figure C.1.14.: Visualisation of step **(3)(b)(i)** with $k_1 = 0$ of the algorithm

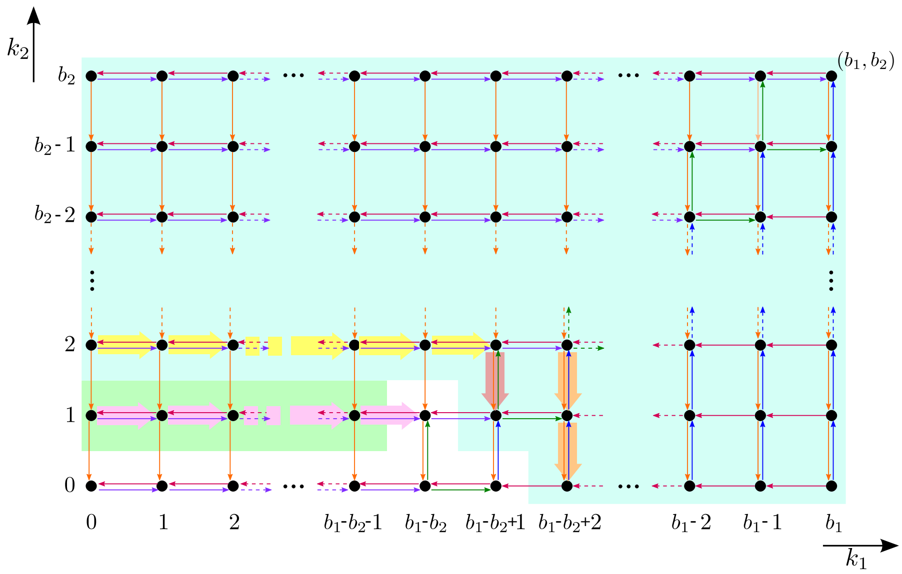
In the next steps, where $\underline{0 < k_1 < b_1 - b_2}$
The GBE of state $(k_1, 1)$ is

$$\widetilde{\theta}(k_1,1)\Big(\lambda_1 \cdot \underbrace{1_{\{k_1>0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{1>0\}}}_{=1} + \underbrace{\sum_{i\in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$
$$= \widetilde{\theta}(k_1,2) \cdot \lambda_2 \cdot \underbrace{1_{\{1<b_2\}}}_{=1}$$
$$+\widetilde{\theta}(k_1+1,1) \cdot \lambda_1 \cdot \underbrace{1_{\{k_1<b_1\}}}_{=1}$$
$$+\widetilde{\theta}(k_1-1,1) \cdot \nu \cdot \underbrace{p_1(\mathbf{k}-\mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{k_1>0\}}}_{=1}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 < b_1 - b_2$ and $k_2 = 1$, we have $b_1 - (k_1 - 1) > b_2 - 1 = b_2 - k_2$, which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$.
- $\widetilde{\theta}(k_1, 1)$ is known for $k_1 > 0$ as a function of κ from the previous steps because in step **(3)(b)**, the loop of k_1 goes from 0 to $b_1 - b_2 + 1$.
  More precisely, in step **(3)(b)(i)**, the GBE of state $(\underline{k_1 - 1}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, 1)$.
- $\widetilde{\theta}(k_1, 2)$ is known for $b_2 > 1$ from the previous for-loop independent of κ because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $b_2 = 2$, because in step **(1)(c)**, the GBE of state $(\underline{k_1 - 1}, b_2) = (\underline{k_1 - 1}, 2)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, b_2) = \widetilde{\theta}(\underline{k_1}, 2)$.
  - if $b_2 > 2$, because in step **(2)(b)(i)**, the GBE of state $(\underline{k_1 - 1}, 2)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, 2)$.
- $\widetilde{\theta}(k_1 - 1, 1)$ is known for $k_1 > 0$ as a function of κ from the previous steps.
  More precisely,
  - if $k_1 = 1$, because $\widetilde{\theta}(k_1 - 1, 1) = \widetilde{\theta}(0, 1)$ was set to κ in step **(3)(a)**,
  - if $k_1 > 1$, because in step **(3)(b)**, the loop of k_1 goes from 0 to $b_1 - b_2 + 1$.
    In step **(3)(b)(i)**, the GBE of state $(\underline{k_1 - 2}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1 - 1}, 1)$.
- Therefore, this GBE can be used to solve for $\widetilde{\theta}(k_1 + 1, 1)$ as a function of κ.



(a) Step **(3)(b)(i)** with $k_1 = 1$



(b) Step **(3)(b)(i)** with $k_1 = b_1 - b_2 - 1$

Figure C.1.15.: Visualisation of step **(3)(b)(i)** of the algorithm

(3)(b)(ii) If $\underline{k_1 = b_1 - b_2}$.
The GBE of state $(b_1 - b_2, 1)$ is

$$\widetilde{\theta}(b_1 - b_2, 1)\Big(\lambda_1 \cdot 1_{\{b_1 - b_2 > 0\}} + \lambda_2 \cdot \underbrace{1_{\{1>0\}}}_{=1} + \underbrace{\sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}(b_1 - b_2, 2) \cdot \lambda_2 \cdot \underbrace{1_{\{1<b_2\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - b_2 + 1, 1) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1 - b_2 < b_1\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - b_2 - 1, 1) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot 1_{\{b_1 - b_2 > 0\}}$$

$$+\widetilde{\theta}(b_1 - b_2, 0) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1/2} \cdot \underbrace{1_{\{1>0\}}}_{=1}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - b_2$ and $k_2 = 1$,
  we have $b_1 - (k_1 - 1) = b_1 - (b_1 - b_2 - 1) = b_2 + 1 > b_2 - 1 = b_2 - k_2$,
  which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{1\}$.
- $p_2(\mathbf{k} - \mathbf{e}_2) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - b_2$ and $k_2 = 1$,
  we have $b_2 - (k_2 - 1) = b_2 - (1 - 1) = b_2 = b_1 - (b_1 - b_2) = b_1 - k_1$,
  which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{1, 2\}$.
- $\widetilde{\theta}(b_1 - b_2, 1)$ is known as a function of κ from the previous steps. More precisely,
  - if $b_1 - b_2 = 0$, because $\widetilde{\theta}(b_1 - b_2, 1) = \widetilde{\theta}(0, 1)$ was set to κ in step **(3)(a)**,
  - if $b_1 - b_2 > 0$, because in step **(3)(b)**, the loop of k_1 goes from 0 to $b_1 - b_2 + 1$. In step **(3)(b)(i)**, the GBE of state $(\underline{b_1 - b_2 - 1}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2}, 1)$.
- $\widetilde{\theta}(b_1 - b_2, 2)$ is known for $b_2 > 1$ from the previous for-loop independent of κ because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $b_2 = 2$, because in step **(1)(c)**, the GBE of state $(\underline{b_1 - b_2 - 1}, b_2) = (\underline{b_1 - 3}, 2)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2}, b_2) = \widetilde{\theta}(\underline{b_1 - 2}, 2)$.
  - if $b_2 > 2$, because in step **(2)(b)(i)**, the GBE of state $(\underline{b_1 - b_2 - 1}, 2)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2}, 2)$.
- $\widetilde{\theta}(b_1 - b_2 + 1, 1)$ is known independent of κ from the previous for-loop because the for-loop of k_2 goes from b_2 to 1.
  More precisely,
  - if $b_2 = 2$, because in step **(1)(e)**, the GBE of state $(b_1 - 1, b_2)$ was used to find an expression for $\widetilde{\theta}(b_1 - b_2 + 1, 1) = \widetilde{\theta}(b_1 - 1, b_2 - 1)$,

- if $b_2 > 2$, because in step **(2)(b)(ii)**, the GBE of state
  $(b_1 - (b_2 - 2) - 1, \underline{2}) = (b_1 - b_2 + 1, \underline{2})$
  was used to find an expression for
  $\widetilde{\theta}\,(b_1 - (b_2 - 2) - 1, \underline{1}) = \widetilde{\theta}\,(b_1 - b_2 + 1, \underline{1})$.

- $\widetilde{\theta}\,(b_1 - b_2 - 1, 1)$ is known for $b_1 - b_2 > 0$ as a function of κ from a previous step because in step **(3)(b),** the loop of k_1 goes from 0 to $b_1 - b_2 + 1$. More precisely,
  - if $b_1 - b_2 - 1 = 0$, then $\widetilde{\theta}\,(b_1 - b_2 - 1, 1) = \widetilde{\theta}(0, 1)$ was set to κ in step **(3)(a)**,
  - if $b_1 - b_2 - 1 > 0$, in step **(3)(b)(i)**, the GBE of state $(\underline{b_1 - b_2 - 2}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2 - 1}, 1)$.

- Therefore, this GBE can be used to solve for $\widetilde{\theta}\,(b_1 - b_2, 0)$ as a function of κ.

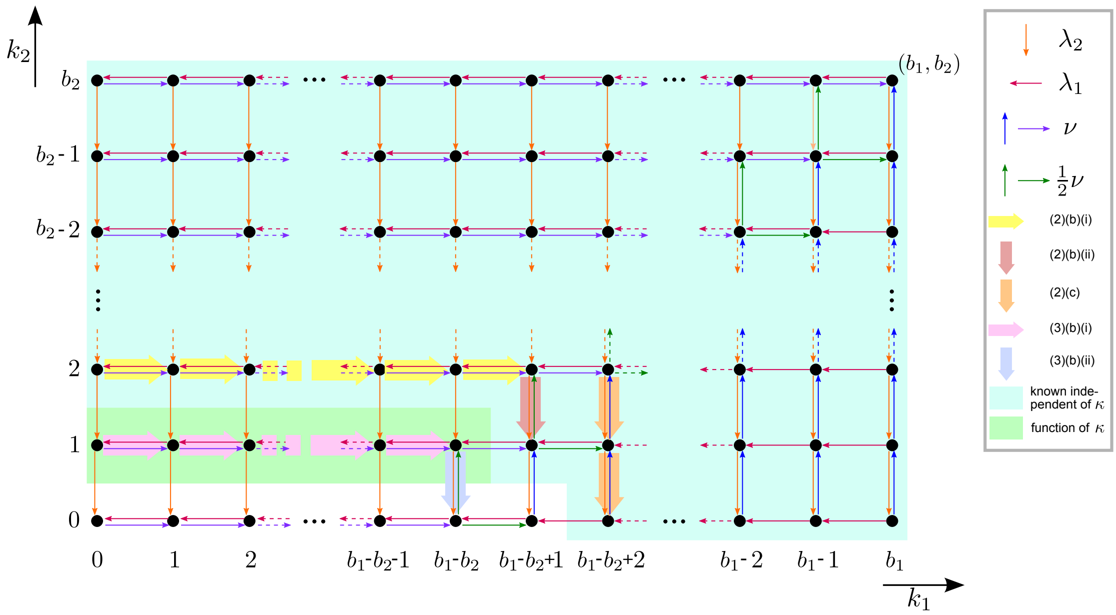


Figure C.1.16.: Visualisation of step **(3)(b)(ii)** with $k_1 = b_1 - b_2$ of the algorithm

If $\underline{k_1 = b_1 - b_2 + 1}$.
The GBE of state $(b_1 - b_2 + 1, 1)$ is

$$\widetilde{\theta}\,(b_1 - b_2 + 1, 1)\Big(\lambda_1 \cdot \underbrace{1_{\{b_1 - b_2 + 1 > 0\}}}_{=1} + \lambda_2 \cdot \underbrace{1_{\{1>0\}}}_{=1} + \underbrace{\sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}\,(b_1 - b_2 + 1, 2) \cdot \lambda_2 \cdot \underbrace{1_{\{1 < b_2\}}}_{=1}$$

$$+ \widetilde{\theta}\,(b_1 - b_2 + 2, 1) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1 - b_2 + 1 < b_1\}}}_{=1}$$

$$+ \widetilde{\theta}\,(b_1 - b_2, 1) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{b_1 - b_2 + 1 > 0\}}}_{=1}$$

$$+ \widetilde{\theta}\,(b_1 - b_2 + 1, 0) \cdot \nu \cdot \underbrace{p_2(\mathbf{k} - \mathbf{e}_2)}_{=1} \cdot \underbrace{1_{\{1>0\}}}_{=1}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - b_2 + 1$ and $k_2 = 1$,
  we have $b_1 - (k_1 - 1) = b_1 - (b_1 - b_2 + 1 - 1) = b_2 > b_2 - 1 = b_2 - k_2$,
  which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{1\}$.

- $p_2(\mathbf{k} - \mathbf{e}_2) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - b_2 + 1$ and $k_2 = 1$,
  we have $b_1 - k_1 = b_1 - (b_1 - b_2 + 1) = b_2 - 1 < b_2 = b_2 - (k_2 - 1)$,
  which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{2\}$.

- $\widetilde{\theta}\,(b_1 - b_2 + 1, 1)$ is known independent of κ from the previous for-loop.
  More precisely,
  - if $b_2 = 2$, because in step **(1)(e)**, the GBE of state $(b_1 - 1, b_2)$ was used to find an expression for $\widetilde{\theta}\,(b_1 - b_2 + 1, 1) = \widetilde{\theta}\,(b_1 - 1, b_2 - 1)$,
  - if $b_2 > 2$, because the for-loop of k_2 goes from b_2 to 1.
    More precisely, because in step **(2)(b)(ii)**, the GBE of state $(b_1 - (b_2 - 2) - 1, \underline{2}) = (b_1 - b_2 + 1, \underline{2})$
    was used to find an expression for
    $\widetilde{\theta}\,(b_1 - (b_2 - 2) - 1, \underline{1}) = \widetilde{\theta}\,(b_1 - b_2 + 1, \underline{1})$.

- $\widetilde{\theta}\,(b_1 - b_2 + 1, 2)$ is known for $b_2 > 1$ from the previous for-loop independent of κ because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $b_2 = 2$, because in step **(1)(c)**, the GBE of state $(\underline{b_1 - 2}, 2)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - 1}, 2)$.
  - if $b_2 > 2$, because in step **(2)(b)(i)**, the GBE of state $(\underline{b_1 - b_2}, b_2)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2 + 1}, b_2)$.

- $\widetilde{\theta}\,(b_1 - b_2 + 2, 1)$ is known independent of κ from the previous for-loop because the loop of k_2 goes from b_2 to 1. More precisely,
  - if $b_2 = 2$, $\widetilde{\theta}\,(b_1 - b_2 + 2, 1) = \widetilde{\theta}\,(b_1, 1)$ is known independent of κ, because in step **(1)(b)** the GBE of state $(b_1, \underline{0})$ was used to find an expression for $\widetilde{\theta}(b_1, \underline{1})$,
  - if $b_2 > 2$, is known independent of κ because the loop of k_2 goes from b_2 to 1. More precisely, in step **(2)(c)** the GBE of state $(b_1 - b_2 + 2, \underline{2})$ was used to find an expression for $\widetilde{\theta}(b_1 - b_2 + 2, \underline{1})$.
- $\widetilde{\theta}\,(b_1 - b_2, 1)$ is known for $b_1 - b_2 + 1 > 0$ as a function of κ from the previous steps. More precisely,
  - if $b_1 - b_2 = 0$, then $\widetilde{\theta}(0, 1)$ was set to κ in step **(3)(a)**,
  - if $b_1 - b_2 > 0$, because in step **(3)(b)**, the loop of k_1 goes from 0 to $b_1 - b_2 + 1$. In step **(3)(b)(i)**, the GBE of state $(\underline{b_1 - b_2 - 1}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2}, 1)$.
- Therefore, this GBE can be used to solve for $\widetilde{\theta}\,(b_1 - b_2 + 1, 0)$ as a function of κ.

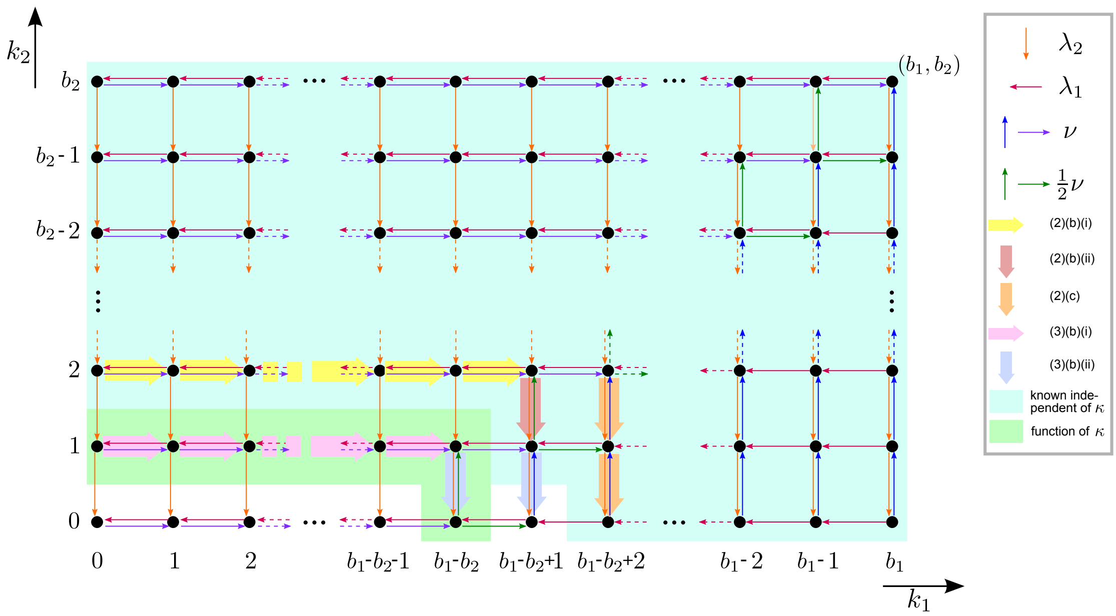


Figure C.1.17.: Visualisation of step **(3)(b)(ii)** with $k_1 = b_1 - b_2 + 1$ of the algorithm

(3)(c) If $b_1 > b_2$, it starts with $\underline{k_1 = b_1 - b_2}$.
The GBE of the state $(b_1 - b_2, 0)$ is

$$\begin{aligned}
&\widetilde{\theta}(b_1 - b_2, 0)\Big(\lambda_1 \cdot 1_{\{b_1 - b_2 > 0\}} + \underbrace{\sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}}}_{=\nu}\Big) \\
= &\widetilde{\theta}(b_1 - b_2, 1) \cdot \lambda_2 \cdot \underbrace{1_{\{0 < b_2\}}}_{=1} \\
&+ \widetilde{\theta}(b_1 - b_2 + 1, 0) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1 - b_2 < b_1\}}}_{=1} \\
&+ \widetilde{\theta}(b_1 - b_2 - 1, 0) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{b_1 - b_2 > 0\}}}_{=1} .
\end{aligned}$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  due to $k_1 = b_1 - b_2$ and $k_2 = 0$,
  we have $b_1 - (k_1 - 1) = b_1 - (b_1 - b_2 - 1) = b_2 + 1 > b_2 = b_2 - k_2$,
  which leads to $\arg\max_{j \in \overline{J}}(b_j - k_j) = \{1\}$.

- $\widetilde{\theta}(b_1 - b_2, 0)$ is known as a function of κ from the previous steps because in step **(3)(b)(ii)** the GBE of state $(b_1 - b_2, \underline{1})$ was used to find an expression for $\widetilde{\theta}(b_1 - b_2, \underline{0})$.

- $\widetilde{\theta}(b_1 - b_2, 1)$ is known for $b_1 - b_2 > 0$ as a function of κ from the previous steps because in step **(3)(b)**, the loop of k_1 goes from 0 to $b_1 - b_2 + 1$. More precisely, in step **(3)(b)(i)**, the GBE of state $(\underline{b_1 - b_2 - 1}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{b_1 - b_2}, 1)$.

- $\widetilde{\theta}(b_1 - b_2 + 1, 0)$ is known for $b_1 - b_2 < b_1$ as a function of κ from the previous steps because in step **(3)(b)**, the loop of k_1 goes from 0 to $b_1 - b_2 + 1$.
  More precisely, in step **(3)(b)(ii)**, the GBE of state $(b_1 - b_2 + 1, \underline{1})$ was used to find an expression for $\widetilde{\theta}(b_1 - b_2 + 1, \underline{0})$.

- Therefore, this GBE can be used to solve for $\widetilde{\theta}(b_1 - b_2 - 1, 0)$ as a function of κ.

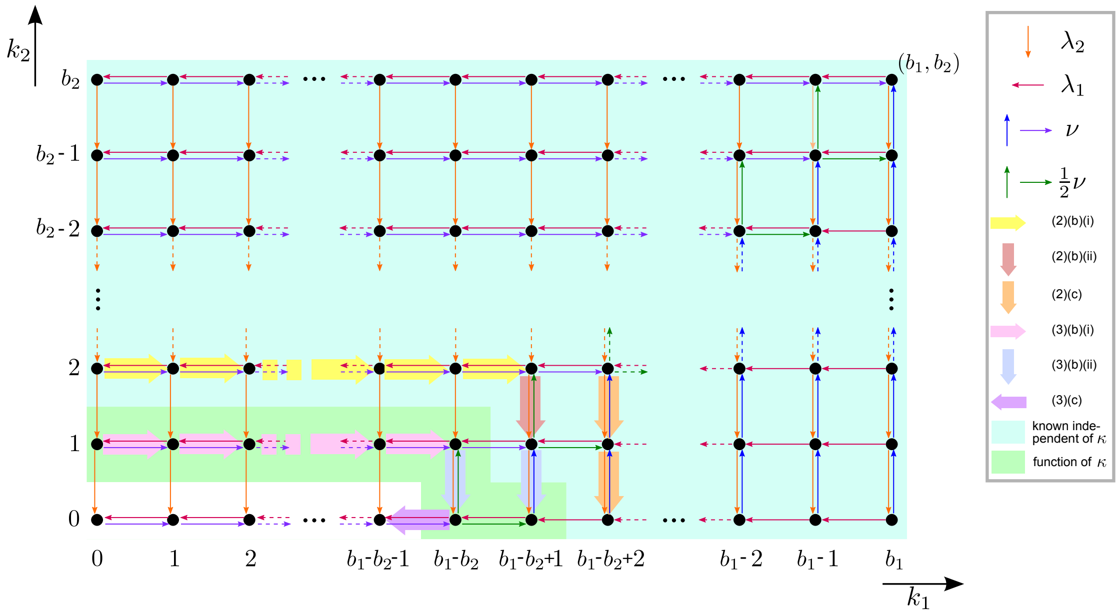


Figure C.1.18.: Visualisation of step **(3)(c)** with $k_1 = b_1 - b_2$ of the algorithm

In the next steps, where $\underline{b_1 - b_2 > k_1 > 0}$.
The GBE of state $(k_1, 0)$ is

$$\underbrace{\widetilde{\theta}\,(k_1,0)}_{=\kappa}\Big(\lambda_1 \cdot \underbrace{1_{\{k_1>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$
$$= \widetilde{\theta}\,(k_1,1) \cdot \lambda_2 \cdot \underbrace{1_{\{0<b_2\}}}_{=1}$$
$$+\widetilde{\theta}\,(k_1+1,0) \cdot \lambda_1 \cdot \underbrace{1_{\{k_1<b_1\}}}_{=1}$$
$$+\widetilde{\theta}\,(k_1-1,0) \cdot \nu \cdot \underbrace{p_1(\mathbf{k}-\mathbf{e}_1)}_{=1} \cdot \underbrace{1_{\{k_1>0\}}}_{=1}.$$

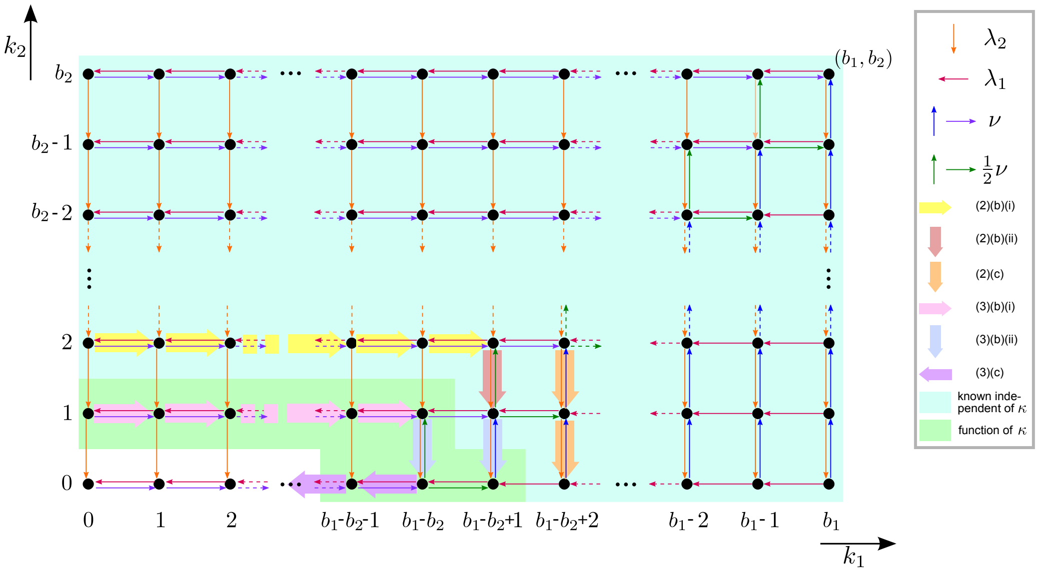
- $p_1(\mathbf{k}-\mathbf{e}_1) = 1$ holds for the following reason:
  Due to $b_1 - b_2 > k_1 > 0$ and $k_2 = 0$,
  we have $b_1 - (k_1 - 1) > b_2 = b_2 - k_2$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1\}$.

- $\widetilde{\theta}\,(k_1, 0)$ is known for $k_1 < b_1 - b_2$ as a function of κ from the previous steps because in step **(3)(c)** the loop of k_1 goes from $b_1 - b_2$ to 1.
  More precisely, in step **(3)(c)**, the GBE of state $(\underline{k_1+1}, 0)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, 0)$.

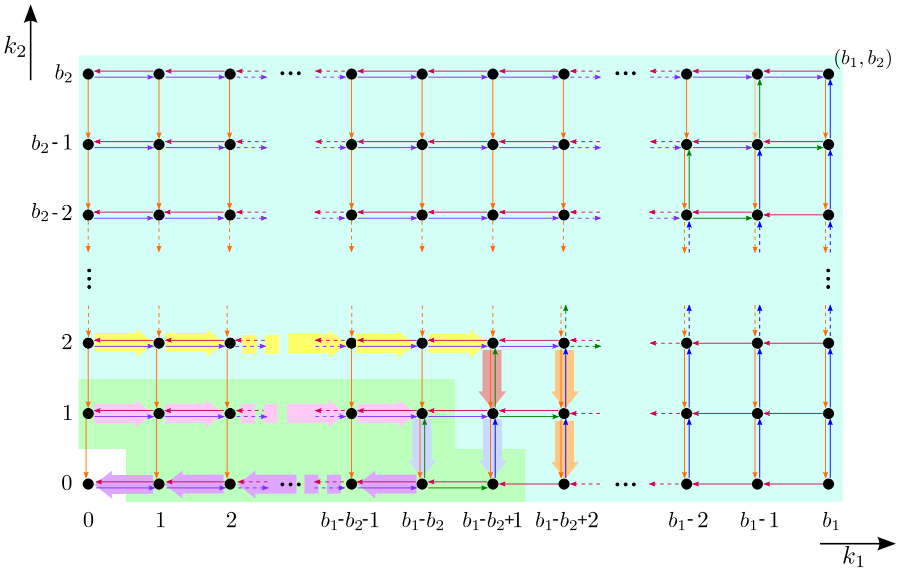
- $\widetilde{\theta}\,(k_1, 1)$ is known for $0 < k_1 < b_1 - b_2$ as a function of κ from the previous steps because in step **(3)(b)(i)**, the GBE of state $(\underline{k_1-1}, 1)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1}, 1)$.

- $\widetilde{\theta}\,(k_1+1, 0)$ is known for $k_1 < b_1$ as a function of κ from the previous steps. More precisely,
  - if $k_1 = b_1 - b_2 - 1$, because $k_1 + 1 = b_1 - b_2$ in step **(3)(b)(ii)**, the GBE of state $(b_1 - b_2, \underline{1}) = (k_1 + 1, \underline{1})$ was used to find an expression for $(b_1 - b_2, \underline{0}) = \widetilde{\theta}(k_1 + 1, \underline{0})$.
  - if $k_1 < b_1 - b_2 - 1$, because in step **(3)(c)** the loop of k_1 goes from $b_1 - b_2$ to 1. More precisely, because $k_1 < b_1 - b_2 - 1$, in step **(3)(c)**, the GBE of state $(\underline{k_1+2}, 0)$ was used to find an expression for $\widetilde{\theta}(\underline{k_1+1}, 0)$.

- Therefore, this GBE can be used to solve for $\widetilde{\theta}\,(k_1 - 1, 0)$ as a function of κ.



(a) Step **(3)(c)** with $k_1 = b_1 - b_2 - 1$



(b) Step **(3)(c)** with $k_1 = 1$

Figure C.1.19.: Visualisation of step **(3)(c)** of the algorithm

(3)(d) The GBE of state $(b_1 - b_2 + 1, 0)$ is

$$\widetilde{\theta}\,(b_1 - b_2 + 1, 0)\Big(\lambda_1 \cdot \underbrace{1_{\{b_1-b_2+1>0\}}}_{=1} + \underbrace{\sum_{i\in\overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i<b_i\}}}_{=\nu}\Big)$$

$$= \widetilde{\theta}\,(b_1 - b_2 + 1), 1) \cdot \lambda_2 \cdot \underbrace{1_{\{0<b_2\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - b_2 + 2, 0) \cdot \lambda_1 \cdot \underbrace{1_{\{b_1-b_2+1<b_1\}}}_{=1}$$

$$+\widetilde{\theta}(b_1 - b_2, 0) \cdot \nu \cdot \underbrace{p_1(\mathbf{k} - \mathbf{e}_1)}_{=1/2} \cdot \underbrace{1_{\{b_1-b_2+1>0\}}}_{=1}.$$

- $p_1(\mathbf{k} - \mathbf{e}_1) = 1$ holds for the following reason:
  Due to $k_1 = b_1 - b_2 + 1$ and $k_2 = 0$,
  we have $b_1 - (k_1 - 1) = b_1 - (b_1 - b_2 + 1 - 1) = b_2 = b_2 - k_2$,
  which leads to $\arg\max_{j\in\overline{J}}(b_j - k_j) = \{1, 2\}$.
- $\widetilde{\theta}\,(b_1 - b_2 + 1, 0)$ is known as a function of κ because in step **(3)(b)(ii)**, the GBE of state $(b_1 - b_2 + 1, \underline{1})$ was used to find an expression for $\widetilde{\theta}(b_1 - b_2 + 1, \underline{0})$.
- $\widetilde{\theta}\,(b_1 - b_2 + 1, 1)$ is known independent of κ from the previous for-loop. More precisely,
  - if $b_2 = 2$, because in step **(1)(e)**, the GBE of state $(b_1 - 1, b_2)$ was used to find an expression for $\widetilde{\theta}\,(b_1 - b_2 + 1, 1) = \widetilde{\theta}\,(b_1 - 1, b_2 - 1)$,
  - if $b_2 > 2$, because the for-loop of k_2 goes from b_2 to 1.
    More precisely, because in step **(2)(b)(ii)**, the GBE of state $(b_1 - (b_2 - 2) - 1, \underline{2}) = (b_1 - b_2 + 1, \underline{2})$
    was used to find an expression for
    $\widetilde{\theta}\,(b_1 - (b_2 - 2) - 1, \underline{1}) = \widetilde{\theta}\,(b_1 - b_2 + 1, \underline{1})$.
- $\widetilde{\theta}\,(b_1 - b_2 + 2, 0)$ is known independent of κ for $b_1 - b_2 + 1 < b_1$. More precisely,
  - if $b_1 = 2$, because $\widetilde{\theta}\,(b_1 - b_2 + 2, 0) = \widetilde{\theta}\,(b_1, 0)$ was set to 1 at the beginning of the algorithm,
  - if $b_1 > 2$, because in step **(2)(c)**, the GBE of state $(b_1 - b_2 + 2, \underline{1})$ was used to find an expression for $\widetilde{\theta}\,(b_1 - b_2 + 2, \underline{0})$.
- $\widetilde{\theta}\,(b_1 - b_2, 0)$ is known as a function of κ because in step **(3)(b)(ii)**, the GBE of state $(b_1 - b_2, \underline{1})$ was used to find an expression for $\widetilde{\theta}(b_1 - b_2, \underline{0})$.

  Hence, everything is known and at least one expression is independent of κ. Thus, we can use this GBE to solve for κ.

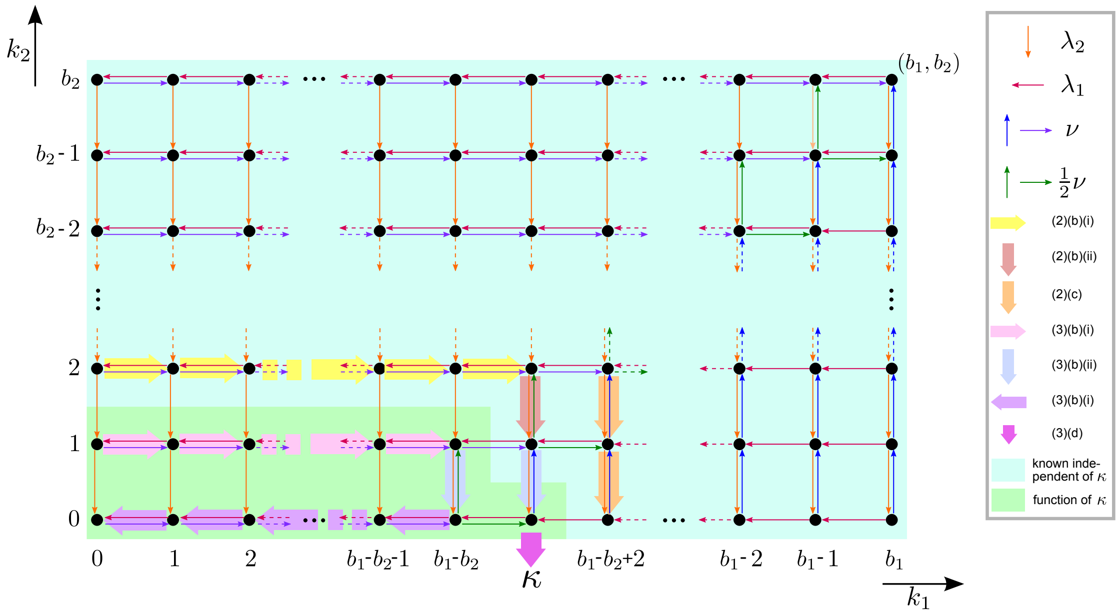


Figure C.1.20.: Visualisation of step **(3)(d)** of the algorithm

Example: Two heterogeneous locations Motivated by the computational example in [CK79, Appendix, pp. 360f.] we consider an example with two locations, base stock levels $b_1 = 4$, $b_2 = 3$ and arrival rates $\lambda_1,\ \lambda_2 > 0$. We recall the notation for this system:

$$\mathbf{k} = \Big(\overbrace{k_1,\ \ k_2,}^{\substack{\text{inventories}\\ \text{at locations}}}\ \overbrace{k_3}^{\text{supplier}} \Big).$$

The state transition diagram is presented in Figure C.1.21.

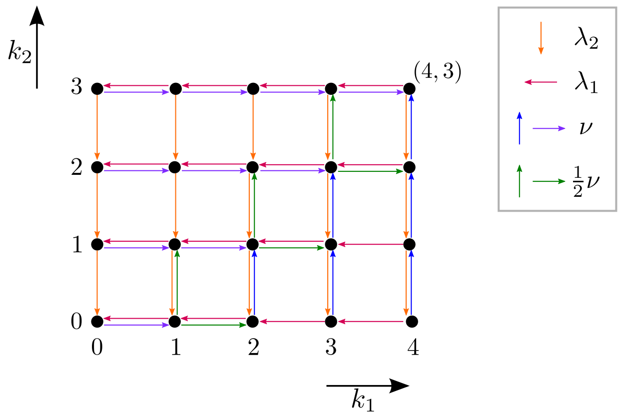


Figure C.1.21.: State transition diagram of a system with two locations and $b_1 = 4$, $b_2 = 3$ and $\lambda_1, \lambda_2 > 0$

Inside the following equations, we use cyan colour for the expressions which are known independent of κ, green colour for the expressions which are known as a function of κ and red colour for the expressions which are unknown.

Example C.1.1. The solution steps for this system using the algorithm follow:

- Set $\widetilde{\theta}(4, 0) = 1$.
- For $k_2 = 3$:

(1)(a) Set $\widetilde{\theta}(0, 3) = \kappa$.

(b) For $\ell = 0$: Use the GBE of state $(4, 0)$ to find an expression for $\widetilde{\theta}(4, 1)$ independent of κ:

$$\widetilde{\theta}(4,0) \cdot (\lambda_1 + \nu) = \widetilde{\theta}(4,1) \cdot \lambda_2,$$

which is equivalent to

$$\widetilde{\theta}(4,1) = \frac{\lambda_1 + \nu}{\lambda_2} =: c^{(4,1)}\left(\lambda_1, \lambda_2, \nu\right).$$

For $\ell = 1$: Use the GBE of state $(4, 1)$ to find an expression for $\widetilde{\theta}(4, 2)$ independent of κ:

$$\widetilde{\theta}(4,1) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(4,2) \cdot \lambda_2,$$

which is equivalent to

$$\widetilde{\theta}(4,2) = \frac{(\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot c^{(4,1)}\left(\lambda_1, \lambda_2, \nu\right) =: c^{(4,2)}\left(\lambda_1, \lambda_2, \nu\right).$$

(c) For $k_1 = 0$: Use the GBE of state $(0, 3)$ to find an expression for $\widetilde{\theta}(1, 3)$ as a function of κ:

$$\widetilde{\theta}(0,3) \cdot (\lambda_2 + \nu) = \widetilde{\theta}(1,3) \cdot \lambda_1,$$

which is equivalent to

$$\widetilde{\theta}(1,3) = \frac{\lambda_2 + \nu}{\lambda_1} \cdot \kappa =: c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

For $k_1 = 1$: Use the GBE of state $(1, 3)$ to find an expression for $\widetilde{\theta}(2, 3)$ as a function of κ:

$$\widetilde{\theta}(1,3) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(2,3) \cdot \lambda_1 + \widetilde{\theta}(0,3) \cdot \nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(2,3) &= \frac{\widetilde{\theta}(1,3) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(0,3) \cdot \nu}{\lambda_1} \\
&= \frac{c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot (\lambda_1 + \lambda_2 + \nu) - \kappa \cdot \nu}{\lambda_1} \\
&= \frac{c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - \nu}{\lambda_1} \cdot \kappa \\
&=: c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.
\end{aligned}$$

For $k_1 = 2$: Use the GBE of state $(2, 3)$ to find an expression for $\widetilde{\theta}(3, 3)$ as a function of κ:

$$\widetilde{\theta}(2,3) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(3,3) \cdot \lambda_1 + \widetilde{\theta}(1,3) \cdot \nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(3,3) &= \frac{\widetilde{\theta}(2,3) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1,3) \cdot \nu}{\lambda_1} \\
&= \frac{c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \nu}{\lambda_1} \\
&= \frac{c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1} \cdot \kappa =: c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.
\end{aligned}$$

(d) Use the GBE of state $(4, 3)$ to find an expression for $\widetilde{\theta}(4, 3)$ as a function of κ:

$$\widetilde{\theta}(4,3) \cdot (\lambda_1 + \lambda_2) = \widetilde{\theta}(3,3) \cdot \nu + \widetilde{\theta}(4,2) \cdot \nu$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(4,3) &= \frac{\widetilde{\theta}(3,3) \cdot \nu + \widetilde{\theta}(4,2) \cdot \nu}{\lambda_1 + \lambda_2} \\
&= \frac{c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \nu + c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1 + \lambda_2} \\
&= \frac{c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1 + \lambda_2} \cdot \kappa + \frac{c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1 + \lambda_2} \\
&=: c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(4,3)}(\lambda_1, \lambda_2, \nu).
\end{aligned}$$

(e) For $k_1 = 3$: Use the GBE of state $(3,3)$
to find an expression for $\widetilde{\theta}(3,2)$ as a function of κ:

$$\widetilde{\theta}(3,3) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(4,3) \cdot \lambda_1 + \widetilde{\theta}(2,3) \cdot \nu + \widetilde{\theta}(3,2) \cdot \frac{1}{2} \cdot \nu$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(3,2) &= \frac{\widetilde{\theta}(3,3) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(4,3) \cdot \lambda_1 - \widetilde{\theta}(2,3) \cdot \nu}{\frac{1}{2} \cdot \nu} \\
&= \frac{c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot (\lambda_1 + \lambda_2 + \nu)}{\frac{1}{2} \cdot \nu} \\
&\quad - \frac{\left(c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(4,3)}(\lambda_1, \lambda_2, \nu)\right) \cdot \lambda_1}{\frac{1}{2} \cdot \nu} \\
&\quad - \frac{c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \nu}{\frac{1}{2} \cdot \nu} \\
&= \frac{c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_1 - c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\frac{1}{2} \cdot \nu} \cdot \kappa \\
&\quad - \frac{c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_1}{\frac{1}{2} \cdot \nu} \\
&=: c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(3,2)}(\lambda_1, \lambda_2, \nu) .
\end{aligned}$$

(f) Use the GBE of state $(4,2)$ to solve for κ.

$$\widetilde{\theta}(4,2) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(4,3) \cdot \lambda_2 + \widetilde{\theta}(3,2) \cdot \frac{1}{2} \cdot \nu + \widetilde{\theta}(4,1) \cdot \nu$$

which is equivalent to

$$\begin{aligned}
& c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) \\
& = \left(c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(4,3)}(\lambda_1, \lambda_2, \nu)\right) \cdot \lambda_2 \\
& \quad + \left(c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(3,2)}(\lambda_1, \lambda_2, \nu)\right) \cdot \frac{1}{2} \cdot \nu + c^{(4,1)}(\lambda_1, \lambda_2, \nu) \cdot \nu
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \lambda_2 + c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \frac{1}{2} \cdot \nu \\
& = c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(4,1)}(\lambda_1, \lambda_2, \nu) \cdot \nu \\
& \quad - c_2^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2} \cdot \nu - c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \quad & \kappa \\
& = \Big[c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(4,1)}(\lambda_1, \lambda_2, \nu) \cdot \nu \\
& \quad - c_2^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2} \cdot \nu - c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_2\Big] \\
& \quad \cdot \frac{1}{c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_2 + c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2} \cdot \nu}
\end{aligned}$$

if the denominator is not equal to zero.

(g) Substitute the value of κ into the equations in the above steps **(1)(a)** and **(1)(c)-(e)**.

▶ For $k_2 = 2$:

(2)(a) Set $\widetilde{\theta}(0,2) = \kappa$.

(b)(i) For $k_1 = 0$: Use the GBE of state $(0,2)$ to find an expression for $\widetilde{\theta}(1,2)$ as a function of κ ($\widetilde{\theta}(0,3)$ is known independent of κ from step **(1)**):

$$\widetilde{\theta}(0,2) \cdot (\lambda_2 + \nu) = \widetilde{\theta}(0,3) \cdot \lambda_2 + \widetilde{\theta}(1,2) \cdot \lambda_1,$$

which is equivalent to

$$\begin{aligned}\widetilde{\theta}(1,2) &= \frac{\widetilde{\theta}(0,3) \cdot \lambda_2 - \widetilde{\theta}(0,2) \cdot (\lambda_2 + \nu)}{\lambda_1} \\ &= \frac{\widetilde{\theta}(0,3) \cdot \lambda_2 - \kappa \cdot (\lambda_2 + \nu)}{\lambda_1} \\ &= \frac{\widetilde{\theta}(0,3) \cdot \lambda_2}{\lambda_1} - \frac{(\lambda_2 + \nu)}{\lambda_1} \cdot \kappa \\ &=: c_1^{(1,2)}(\lambda_1, \lambda_2) + c_2^{(1,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.\end{aligned}$$

For $k_1 = 1$: Use the GBE of state $(1,2)$ to find an expression for $\widetilde{\theta}(2,2)$ as a function of κ ($\widetilde{\theta}(1,3)$ is known independent of κ from step **(1)**):

$$\widetilde{\theta}(1,2) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(2,2) \cdot \lambda_1 + \widetilde{\theta}(1,3) \cdot \lambda_2 + \widetilde{\theta}(0,2) \cdot \nu,$$

which is equivalent to

$$\begin{aligned}\widetilde{\theta}(2,2) &= \frac{\widetilde{\theta}(1,2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1,3) \cdot \lambda_2 - \widetilde{\theta}(0,2) \cdot \nu}{\lambda_1} \\ &= \frac{\left(c_1^{(1,2)}(\lambda_1, \lambda_2) + c_2^{(1,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa\right) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1,3) \cdot \lambda_2 - \kappa \cdot \nu}{\lambda_1} \\ &= \frac{c_1^{(1,2)}(\lambda_1, \lambda_2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1,3) \cdot \lambda_2}{\lambda_1} \\ &\quad + \frac{\left[c_2^{(1,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - \nu\right]}{\lambda_1} \cdot \kappa \\ &=: c_1^{(2,2)}(\lambda_1, \lambda_2, \nu) + c_2^{(2,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.\end{aligned}$$

(b)(ii) For $k_1 = 2$: Use the GBE of state $(2,2)$ to find an expression for $\widetilde{\theta}(2,1)$ as a function of κ ($\widetilde{\theta}(3,2)$ and $\widetilde{\theta}(2,3)$ are known independent of κ from step **(1)**):

$$\widetilde{\theta}(2,2)\cdot(\lambda_1+\lambda_2+\nu) = \widetilde{\theta}(3,2)\cdot\lambda_1 + \widetilde{\theta}(2,3)\cdot\lambda_2 + \widetilde{\theta}(1,2)\cdot\nu + \widetilde{\theta}(2,1)\cdot\frac{1}{2}\nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(2,1) &= \frac{\widetilde{\theta}(2,2)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(3,2)\cdot\lambda_1 - \widetilde{\theta}(2,3)\cdot\lambda_2 - \widetilde{\theta}(1,2)\cdot\nu}{\frac{1}{2}\nu}\\
&= \frac{\left(c_1^{(2,2)}(\lambda_1,\lambda_2,\nu) + c_2^{(2,2)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\cdot(\lambda_1+\lambda_2+\nu)}{\frac{1}{2}\nu}\\
&\quad - \frac{\widetilde{\theta}(3,2)\cdot\lambda_1 + \widetilde{\theta}(2,3)\cdot\lambda_2 + \left(c_1^{(1,2)}(\lambda_1,\lambda_2) + c_2^{(1,2)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\cdot\nu}{\frac{1}{2}\nu}\\
&= \frac{c_1^{(2,2)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(3,2)\cdot\lambda_1 - \widetilde{\theta}(2,3)\cdot\lambda_2 - c_1^{(1,2)}(\lambda_1,\lambda_2)\cdot\nu}{\frac{1}{2}\nu}\\
&\quad + \frac{c_2^{(2,2)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\lambda_2+\nu) - c_2^{(1,2)}(\lambda_1,\lambda_2,\nu)\cdot\nu}{\frac{1}{2}\nu}\cdot\kappa\\
&=: c_1^{(2,1)}(\lambda_1,\lambda_2,\nu) + c_2^{(2,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa.
\end{aligned}$$

(c) For $\ell = 2$: Use the GBE of state $(3,2)$ to find an expression for $\widetilde{\theta}(3,1)$ as a function of κ ($\widetilde{\theta}(3,2)$, $\widetilde{\theta}(4,2)$ and $\widetilde{\theta}(3,3)$ are known independent of κ from step **(1)**):

$$\widetilde{\theta}(3,2)\cdot(\lambda_1+\lambda_2+\nu) = \widetilde{\theta}(4,2)\cdot\lambda_1 + \widetilde{\theta}(3,3)\cdot\lambda_2 + \widetilde{\theta}(2,2)\cdot\nu + \widetilde{\theta}(3,1)\cdot\nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(3,1) &= \frac{\widetilde{\theta}(3,2)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(4,2)\cdot\lambda_1 - \widetilde{\theta}(3,3)\cdot\lambda_2 - \widetilde{\theta}(2,2)\cdot\nu}{\nu}\\
&= \frac{\widetilde{\theta}(3,2)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(4,2)\cdot\lambda_1 - \widetilde{\theta}(3,3)\cdot\lambda_2}{\nu}\\
&\quad - \left(c_1^{(2,2)}(\lambda_1,\lambda_2,\nu) + c_2^{(2,2)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\\
&= \frac{\widetilde{\theta}(3,2)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(4,2)\cdot\lambda_1 - \widetilde{\theta}(3,3)\cdot\lambda_2 - c_1^{(2,2)}(\lambda_1,\lambda_2,\nu)\cdot\nu}{\nu}\\
&\quad - c_2^{(2,2)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\\
&=: c_1^{(3,1)}(\lambda_1,\lambda_2,\nu) + c_2^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa.
\end{aligned}$$

For $\ell = 1$: Use the GBE of state $(3,1)$ to find an expression for $\widetilde{\theta}(3,0)$ as a function of κ ($\widetilde{\theta}(4,1)$ and $\widetilde{\theta}(3,2)$ are known independent of κ from step **(1)**):

$$\widetilde{\theta}(3,1)\cdot(\lambda_1+\lambda_2+\nu) = \widetilde{\theta}(4,1)\cdot\lambda_1 + \widetilde{\theta}(3,2)\cdot\lambda_2 + \widetilde{\theta}(2,1)\cdot\frac{1}{2}\nu + \widetilde{\theta}(3,0)\cdot\nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(3,0) &= \frac{\widetilde{\theta}(3,1)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(4,1)\cdot\lambda_1 - \widetilde{\theta}(3,2)\cdot\lambda_2 - \widetilde{\theta}(2,1)\cdot\frac{1}{2}\nu}{\nu} \\
&= \frac{\left(c_1^{(3,1)}(\lambda_1,\lambda_2,\nu) + c_2^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\cdot(\lambda_1+\lambda_2+\nu)}{\nu} \\
&\quad - \frac{\widetilde{\theta}(4,1)\cdot\lambda_1 + \widetilde{\theta}(3,2)\cdot\lambda_2}{\nu} \\
&\quad - \frac{\left(c_1^{(2,1)}(\lambda_1,\lambda_2,\nu) + c_2^{(2,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\cdot\frac{1}{2}\nu}{\nu} \\
&= \frac{c_1^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\lambda_2+\nu) - \widetilde{\theta}(4,1)\cdot\lambda_1 - \widetilde{\theta}(3,2)\cdot\lambda_2 - c_1^{(2,1)}(\lambda_1,\lambda_2,\nu)\frac{1}{2}\nu}{\nu} \\
&\quad + \frac{c_2^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa(\lambda_1+\lambda_2+\nu) - c_2^{(2,1)}(\lambda_1,\lambda_2,\nu)\cdot\frac{1}{2}\nu}{\nu}\cdot\kappa \\
&=: c_1^{(3,0)}(\lambda_1,\lambda_2,\nu) + c_2^{(3,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa.
\end{aligned}$$

(d) Use the GBE of state $(3,0)$ to solve for κ
($\widetilde{\theta}(4,0)$ is known independent of κ from step **(1)**):

$$\widetilde{\theta}(3,0)\cdot(\lambda_1+\nu) = \widetilde{\theta}(4,0)\cdot\lambda_1 + \widetilde{\theta}(3,1)\cdot\lambda_2,$$

which is equivalent to

$$\begin{aligned}
&\left(c_1^{(3,0)}(\lambda_1,\lambda_2,\nu) + c_2^{(3,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\cdot(\lambda_1+\nu) \\
&= \widetilde{\theta}(4,0)\cdot\lambda_1 + \left(c_1^{(3,1)}(\lambda_1,\lambda_2,\nu) + c_2^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\right)\cdot\lambda_2 \\
\Leftrightarrow\quad & c_2^{(3,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\cdot(\lambda_1+\nu) - c_2^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\kappa\cdot\lambda_2 \\
&= \widetilde{\theta}(4,0)\cdot\lambda_1 + c_1^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\lambda_2 - c_1^{(3,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu) \\
\Leftrightarrow\quad & \kappa \\
&= \left[\widetilde{\theta}(4,0)\cdot\lambda_1 + c_1^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\lambda_2 - c_1^{(3,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu)\right] \\
&\quad \cdot\frac{1}{c_2^{(3,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu) - c_2^{(3,1)}(\lambda_1,\lambda_2,\nu)\cdot\lambda_2}
\end{aligned}$$

if the denominator is not equal to zero.

(3)(a) Set $\widetilde{\theta}(0,1) = \kappa$.

(b)(i) For $k_1 = 0$: Use the GBE of state $(0,1)$
to find an expression for $\widetilde{\theta}(1,1)$ as a function of κ
($\widetilde{\theta}(0,2)$ is known independent of κ from step **(2)**):

$$\widetilde{\theta}(0,1)\cdot(\lambda_2+\nu) = \widetilde{\theta}(0,2)\cdot\lambda_2 + \widetilde{\theta}(1,1)\cdot\lambda_1,$$

which is equivalent to

$$\widetilde{\theta}(1,1) = \frac{(\lambda_2+\nu)}{\lambda_1}\cdot\kappa - \frac{\widetilde{\theta}(0,2)\cdot\lambda_2}{\lambda_1} =: c_1^{(1,1)}(\lambda_1,\lambda_2)\cdot\kappa + c_2^{(1,1)}(\lambda_1,\lambda_2).$$

(b)(ii) For $k_1 = 1$: Use the GBE of state $(1,1)$
to find an expression for $\widetilde{\theta}(1,0)$ as a function of κ
($\widetilde{\theta}(2,1)$ and $\widetilde{\theta}(1,2)$ are known independent of κ from step **(2)**):

$$\widetilde{\theta}(1,1) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(2,1) \cdot \lambda_1 + \widetilde{\theta}(1,2) \cdot \lambda_2 + \widetilde{\theta}(0,1) \cdot \nu + \widetilde{\theta}(1,0) \cdot \frac{1}{2}\nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(1,0) &= \frac{\widetilde{\theta}(1,1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(2,1) \cdot \lambda_1 - \widetilde{\theta}(1,2) \cdot \lambda_2 - \widetilde{\theta}(0,1) \cdot \nu}{\frac{1}{2}\nu} \\
&= \frac{\left(c_1^{(1,1)}(\lambda_1,\lambda_2) \cdot \kappa + c_2^{(1,1)}(\lambda_1,\lambda_2)\right) \cdot (\lambda_1 + \lambda_2 + \nu)}{\frac{1}{2}\nu} \\
&\quad - \frac{\widetilde{\theta}(2,1) \cdot \lambda_1 + \widetilde{\theta}(1,2) \cdot \lambda_2 + \kappa \cdot \nu}{\frac{1}{2}\nu} \\
&= \frac{c_1^{(1,1)}(\lambda_1,\lambda_2) \cdot (\lambda_1 + \lambda_2 + \nu) - \nu}{\frac{1}{2}\nu} \cdot \kappa \\
&\quad + \frac{c_2^{(1,1)}(\lambda_1,\lambda_2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(2,1) \cdot \lambda_1 - \widetilde{\theta}(1,2) \cdot \lambda_2}{\frac{1}{2}\nu} \\
&=: c_1^{(1,0)}(\lambda_1,\lambda_2,\nu) \cdot \kappa + c_2^{(1,0)}(\lambda_1,\lambda_2,\nu).
\end{aligned}$$

For $k_1 = 2$: Use the GBE of state $(2,1)$
to find an expression for $\widetilde{\theta}(2,0)$ as a function of κ
($\widetilde{\theta}(2,1)$, $\widetilde{\theta}(3,2)$ and $\widetilde{\theta}(2,2)$ are known independent of κ from steps **(1)** and **(2)**):

$$\widetilde{\theta}(2,1) \cdot (\lambda_1 + \lambda_2 + \nu) = \widetilde{\theta}(3,2) \cdot \lambda_1 + \widetilde{\theta}(2,2) \cdot \lambda_2 + \widetilde{\theta}(2,0) \cdot \nu + \widetilde{\theta}(1,1) \cdot \nu,$$

which is equivalent to

$$\begin{aligned}
\widetilde{\theta}(2,0) &= \frac{\widetilde{\theta}(2,1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3,2) \cdot \lambda_1 - \widetilde{\theta}(2,2) \cdot \lambda_2 - \widetilde{\theta}(1,1) \cdot \nu}{\nu} \\
&= \frac{\widetilde{\theta}(2,1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3,2) \cdot \lambda_1 - \widetilde{\theta}(2,2) \cdot \lambda_2}{\nu} \\
&\quad - \frac{\left(c_1^{(1,1)}(\lambda_1,\lambda_2) \cdot \kappa + c_2^{(1,1)}(\lambda_1,\lambda_2)\right) \cdot \nu}{\nu} \\
&= -\frac{c_1^{(1,1)}(\lambda_1,\lambda_2) \cdot \nu}{\nu} \cdot \kappa \\
&\quad + \frac{\widetilde{\theta}(2,1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3,2) \cdot \lambda_1 - \widetilde{\theta}(2,2) \cdot \lambda_2 - c_2^{(1,1)}(\lambda_1,\lambda_2) \cdot \nu}{\nu} \\
&=: c_1^{(2,0)}(\lambda_1,\lambda_2,\nu) \cdot \kappa + c_2^{(2,0)}(\lambda_1,\lambda_2,\nu).
\end{aligned}$$

(c) For $k_1 = 1$: Use the GBE of state $(1,0)$
to find an expression for $\widetilde{\theta}(0,0)$ as a function of κ:

$$\widetilde{\theta}(1,0) \cdot (\lambda_1 + \nu) = \widetilde{\theta}(2,0) \cdot \lambda_1 + \widetilde{\theta}(1,1) \cdot \lambda_2 + \widetilde{\theta}(0,0) \cdot \nu,$$

which is equivalent to

$$
\begin{aligned}
\widetilde{\theta}(0,0) &= \frac{\widetilde{\theta}(1,0)\cdot(\lambda_1+\nu) - \widetilde{\theta}(2,0)\cdot\lambda_1 - \widetilde{\theta}(1,1)\cdot\lambda_2}{\nu} \\
&= \frac{\left(c_1^{(1,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa + c_2^{(1,0)}(\lambda_1,\lambda_2,\nu)\right)\cdot(\lambda_1+\nu)}{\nu} \\
&\quad - \frac{\left(c_1^{(2,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa + c_2^{(2,0)}(\lambda_1,\lambda_2,\nu)\right)\cdot\lambda_1}{\nu} \\
&\quad - \frac{\left(c_1^{(1,1)}(\lambda_1,\lambda_2)\cdot\kappa + c_2^{(1,1)}(\lambda_1,\lambda_2)\right)\cdot\lambda_2}{\nu} \\
&= \frac{c_1^{(1,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu) - c_1^{(2,0)}(\lambda_1,\lambda_2,\nu)\cdot\lambda_1 - c_1^{(1,1)}(\lambda_1,\lambda_2)\cdot\lambda_2}{\nu}\cdot\kappa \\
&\quad + \frac{c_2^{(1,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu) - c_2^{(2,0)}(\lambda_1,\lambda_2,\nu)\cdot\lambda_1 - c_2^{(1,1)}(\lambda_1,\lambda_2)\cdot\lambda_2}{\nu} \\
&=: c_1^{(0,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa + c_2^{(0,0)}(\lambda_1,\lambda_2,\nu)\,.
\end{aligned}
$$

(d) Use the GBE of state $(2,0)$ to solve for κ ($\widetilde{\theta}(3,0)$, $\widetilde{\theta}(2,1)$ are known independent of κ from steps **(2)**):

$$
\begin{aligned}
&\widetilde{\theta}(2,0)\cdot(\lambda_1+\nu) \\
&= \widetilde{\theta}(3,0)\cdot\lambda_1 + \widetilde{\theta}(2,1)\cdot\lambda_2 + \widetilde{\theta}(1,0)\cdot\frac{1}{2}\nu
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow\quad &\left(c_1^{(2,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa + c_2^{(2,0)}(\lambda_1,\lambda_2,\nu)\right)\cdot(\lambda_1+\nu) \\
&= \widetilde{\theta}(3,0)\cdot\lambda_1 + \widetilde{\theta}(2,1)\cdot\lambda_2 \\
&\quad + \left(c_1^{(1,0)}(\lambda_1,\lambda_2,\nu)\cdot\kappa + c_2^{(1,0)}(\lambda_1,\lambda_2,\nu)\right)\cdot\frac{1}{2}\nu
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow\quad &\kappa\cdot\left(c_1^{(2,0)}(\lambda_1,\lambda_2,\nu) - c_1^{(1,0)}(\lambda_1,\lambda_2,\nu)\right) \\
&= \widetilde{\theta}(3,0)\cdot\lambda_1 + \widetilde{\theta}(2,1)\cdot\lambda_2 - c_2^{(2,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu) + c_2^{(1,0)}(\lambda_1,\lambda_2,\nu)\cdot\frac{1}{2}\nu
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow\quad &\kappa \\
&= \frac{\widetilde{\theta}(3,0)\cdot\lambda_1 + \widetilde{\theta}(2,1)\cdot\lambda_2 - c_2^{(2,0)}(\lambda_1,\lambda_2,\nu)\cdot(\lambda_1+\nu) + c_2^{(1,0)}(\lambda_1,\lambda_2,\nu)\cdot\frac{1}{2}\nu}{c_1^{(2,0)}(\lambda_1,\lambda_2,\nu) - c_1^{(1,0)}(\lambda_1,\lambda_2,\nu)}
\end{aligned}
$$

if the denominator is not equal to zero.

Special case: $b_1 > b_2 = 1$

If we consider a system with base stock levels $b_1 > 1$ and $b_2 = 1$ the algorithm reduces to the following algorithm on the next page. Step **(1)(b)** and step **(2)** are cancelled because $b_2 = 1$. Furthermore, in step **(3)** we can skip directly to step **(3)(c)** and we do not need to set a prenormalised probability to κ. In Figure C.1.22 the steps of the algorithm are visualised.

ALGORITHM ($b_1 \geq b_2$ with $b_1 > 1$, $b_2 = 1$ and $\lambda_1, \lambda_2 > 0$)

- Set $\widetilde{\theta}(b_1, 0) = 1$.

(1)(a) set $\widetilde{\theta}(0, 1) = \kappa$

(c) for $k_1 = 0, \ldots, b_1 - 2$
use the GBE of state $(k_1, 1)$
to find an expression for $\widetilde{\theta}(k_1 + 1, 1)$ as a function of κ

(d) use the GBE of state $(b_1, 1)$
to find an expression for $\widetilde{\theta}(b_1, 1)$ as a function of κ

(e) use the GBE of state $(b_1 - 1, 1)$
to find an expression for $\widetilde{\theta}(b_1 - 1, 0)$ as a function of κ

(f) use the GBE of state $(b_1, 0)$ to solve for κ

(g) substitute the value of κ into the equations from steps **(1)(a)** and **(1)(c)-(e)**

(3)(c) for $k_1 = b_1 - 1, \ldots, 1$,
use the GBE of state $(k_1, 0)$
to find an expression for $\widetilde{\theta}(k_1 - 1, 0)$

- Normalise all $\widetilde{\theta}(k_1, k_2)$ by setting

$$\widetilde{\theta}(k_1, k_2) \leftarrow \frac{\widetilde{\theta}(k_1, k_2)}{\sum_{k_1=0}^{b_1} \sum_{k_2=0}^{b_2} \widetilde{\theta}(k_1, k_2)}.$$

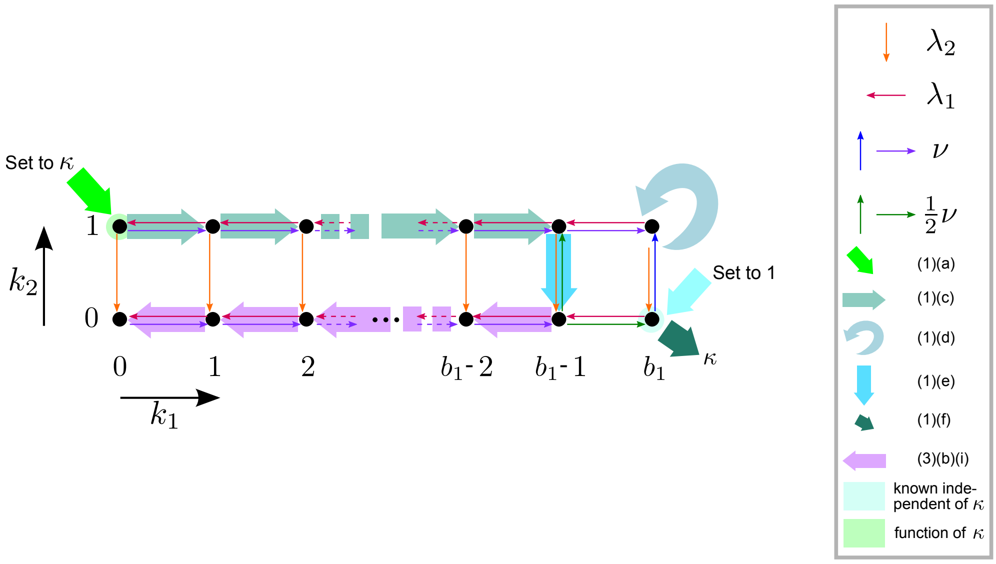


Figure C.1.22.: Visualisation of the steps of the algorithm with $b_1 > b_2 = 1$

Special case: Two homogeneous locations
Our algorithm for two locations can also be applied to obtain $\widetilde{\theta}$ for a system with two homogeneous locations, i.e. where the inventories have identical base stock levels $b_1 = b_2 > 1$ and identical arrival rates $\lambda_1 = \lambda_2 > 0$ (and any service rates $\mu_1, \mu_2 > 0$). The state transition diagram for such a system is presented in Figure C.1.23.

However, the algorithm can be simplified for the case of two homogeneous locations, since in the state transition diagram of a system with two homogeneous locations is symmetric about the diagonal elements (k_1, k_2). Hence, the state transition diagram can be folded to obtain a triangle as shown in Figure C.1.24. Consequently, it holds $\widetilde{\theta}(k_1, k_2) = \widetilde{\theta}(k_2, k_1)$.

Chow and Kohler analyse in [CK77] the performance of homogeneous two-processor distributed computer systems under several dynamic load balancing policies. Their analysis is based on the recursive solution technique and they illustrated the algorithm for a sample system [CK77, Appendix, pp. 51f.]. Their strategy "join the shorter queue without channel transfer" in Chow and Kohler's Model B [CK77, pp. 42f.] is equivalent to our strict load balancing policy.

Additionally, they allow a channel transfer in Model C [CK77, pp. 42f.] and show that under heavy load this strategy can improve the performance (turnaround time) of the two-processor system (cf. [CK77, pp. 47f.]).

We can also extend our model with two locations by a channel transfer and obtain a stationary distribution of product form. Then, the inventories at the locations are connected through a transfer channel. The transfer time of the channel is exponentially distributed with rate $\beta > 0$. If the difference between the number of items in location i and the number of items at location j is equal to or greater than two ($\equiv$ disbalance condition), then the transfer channel initiates a transfer from items from location i to location j. There can be only one transfer at a time and the transfer of an item is discontinued if the disbalance condition changes before the transfer is completed. The state transition diagram can still be folded into a triangle, so that the symmetry property holds. Second, because the channel transfer leads to a further summand on the left side of the GBE (flow into the state). The state transition diagram as well as the folded state transition diagram for such a system with two locations are presented in Figure C.1.25 and Figure C.1.26.

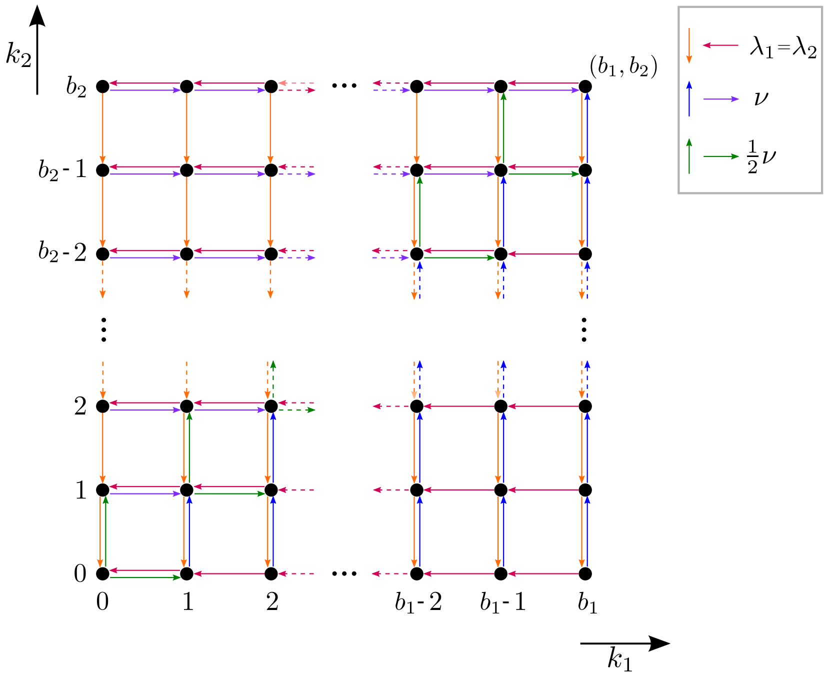


Figure C.1.23.: State transition diagram of a system with two homogeneous locations

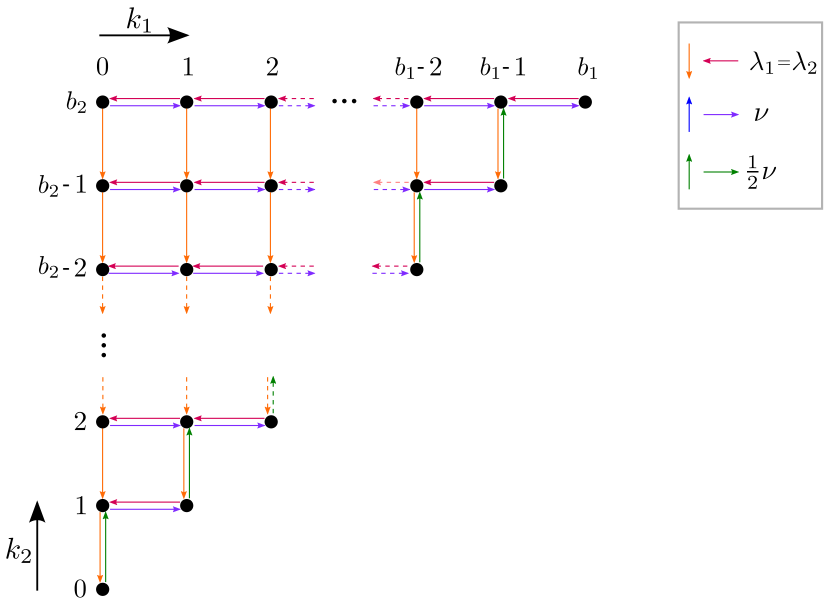


Figure C.1.24.: Folded state transition diagram of a system with two homogeneous locations

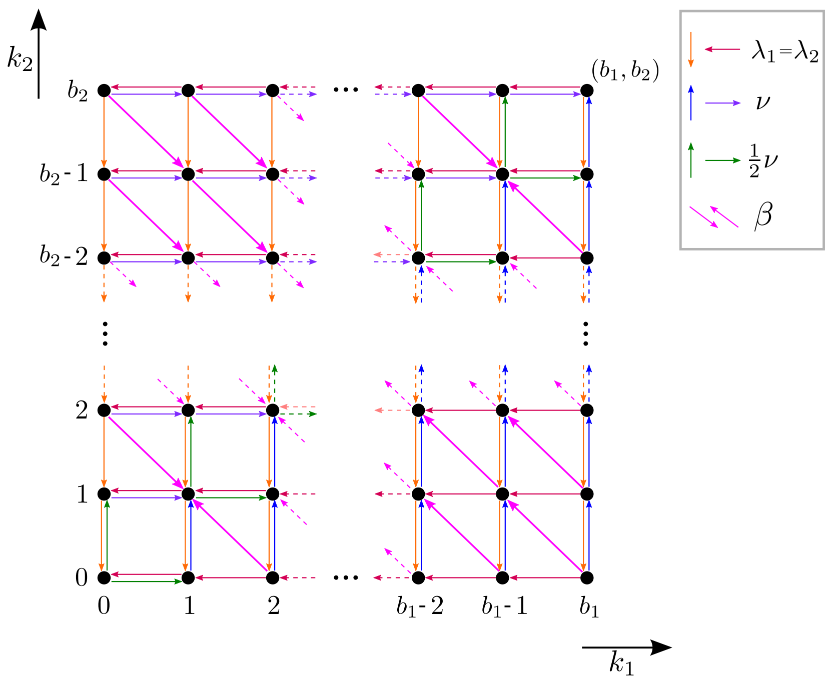


Figure C.1.25.: State transition diagram of a system with two homogeneous locations and transfer channel

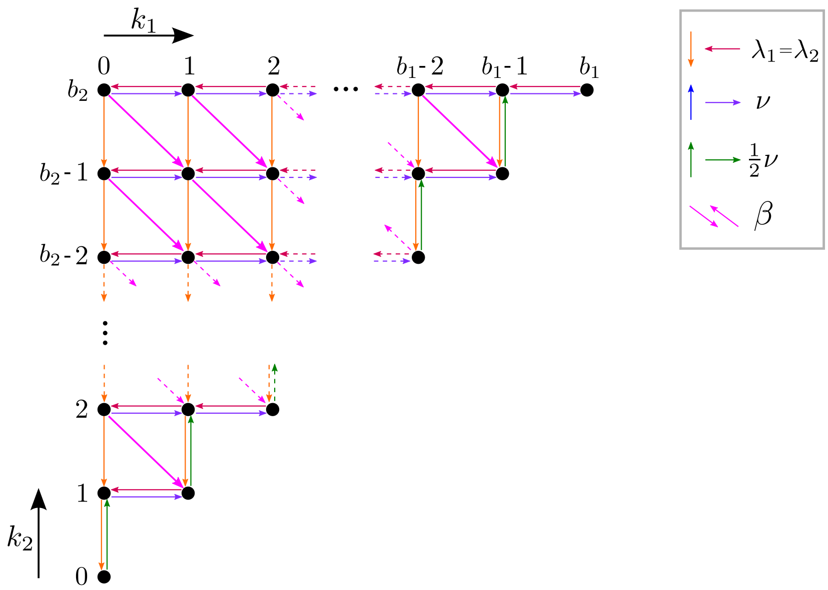


Figure C.1.26.: Folded state transition diagram of a system with two homogeneous locations and transfer channel

D. Appendix to Chapter 4

D.1. Queueing system in a random environment

In this section, we consider a queueing system, which consists of a single server system under a first-come, first-served (FCFS) regime with infinite waiting room, in a random environment, where the queueing system and the environment interact in both directions as in [KD15] and [Kre16, Section 2]. The queueing-environment system of interest is depicted in Figure D.1.1. For a literature review about related research on queueing system in a random environment we refer to [KD15, p. 129] and [Kre16, Section 1].

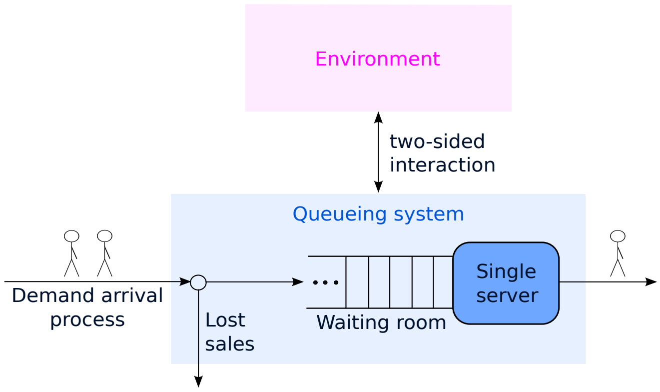


Figure D.1.1.: Queueing-environment system

For the integrated queueing-environment system, we denote by $(Y(t) : t \geq 0)$ the environment process and by $(X(t) : t \geq 0)$ the queue length process. Therefore, by $Z = ((X(t), Y(t)) : t \geq 0)$ we define the joint queueing-environment process of this system.

The state space of Z is $E = \mathbb{N}_0 \times K$, where $\mathbb{N}_0$ denotes the queue length and K is the environment space of the process. We assume $|K| < \infty$. The environment space is partitioned into disjoint components $K := K_W \uplus K_B$. Whenever the environment process enters K_B, the service process is completely "blocked", i.e. service is interrupted and newly arriving customers are lost. Whenever the environment process returns to K_W, the server "works" again and the arrival process is resumed. We assume that the sets K_W and K_B are not empty.

Facilities in the queueing-environment system. The queueing system consists of a single server with infinite waiting room that serves customers on a make-to-order basis under a first-come, first-served (FCFS) regime. Customers arrive one by one at the system according to a Poisson process with queue-length-dependent intensities. If the environment is in state $Y(t) = k \in K_W$ and if there are $X(t) = n \geq 0$ customers present in the queueing system at time $t \geq 0$, customers arrive at the system with intensity $\lambda(n) > 0$.

The queueing-environment system develops over time as follows:

(1) If the environment is in state $Y(t) = k \in K_W$ and if there are $X(t) = n > 0$ customers present in the queueing system (either waiting or in service) at time $t \geq 0$, service is provided to the customer at the head of the line with intensity $\mu(n) > 0$. A customer departs from the system immediately after service and the environment changes with probability $R_n(k, \ell)$ to state $\ell \in K$, independent of the history of the system, given k and n. We consider $R_n = (R_n(k, \ell) : k, \ell \in K)$, $n \in \mathbb{N}$, as a stochastic matrix for the environment driven by the departure process.

(2) If the environment is in state $Y(t) = k \in K_B$ at time $t \geq 0$, no service is provided to customers in the queue and new arriving customers decide not to join the queue and are lost ("lost sales").

(3) Whenever the environment is in state $Y(t) = k \in K$ and if there are $X(t) = n$ customers present at the queueing system (either waiting or in service) at time $t \geq 0$, then the environment changes with rate $v_n(k, \ell) \geq 0$ to state $\ell \in K$, $\ell \neq k$, independent of the history of the system, given k and n. We consider $V_n = (v_n(k, \ell) : k, \ell \in K)$ for all $n \in \mathbb{N}_0$, as a generator, i.e. $\sum_{\ell \in K} v_n(k, \ell) = 0$ for all $k \in K$.

To obtain a Markovian process description of the integrated queueing-environment system via Z we make the usual independence and memoryless assumptions. Then Z is a homogeneous Markov process. We assume henceforth that Z is irreducible and regular.

Remark D.1.1. In [KD15, Section 2], the stochastic matrix for the environment $R = (R(k, \ell) : k, \ell \in K)$ and the generator $V = (v(k, \ell) : k, \ell \in K)$ do not depend on n. In contrast, in (1) and (3) above the changes of the environment's status are coupled with the queue length process.

In [KD15, Section 2], the set K_B may be empty, this means that no interruption of the server and the arrival process occurs.

Krenzler and Daduna prove in [KD15, Section 2] for the case $V_n = V$, $n \in \mathbb{N}$, and $R_n = R$, $n \in \mathbb{N}_0$, a product form steady state distribution of the joint queueing-environment process. The problem of our model is that at present the stationary distribution seems to be out of reach.

Example D.1.2. Special cases of the model in this section are the queueing-inventory systems described in Chapter 4 (server = production system, environment = inventory-replenishment subsystem). The queueing-inventory system is depicted in Figure D.1.2. The corresponding results are presented in Table D.1 on page 348.

(a) The queueing-inventory system with exponentially distributed life times described in Section 4.2.3 fits into the definition of the queueing system in a random environment by setting

$$K = \{(k, b-k) : k \in \{0, 1, \ldots, b\}\},$$

$$K_B = \{(k, b-k) : k = 0\}, \qquad K_W = \{(k, b-k) : k \in \{1, \ldots, b\}\},$$

where K_W indicates for the inventory that there is stock on hand for production, and

$$R_n((0, b), (0, b)) = 1, \qquad R_n((k, b-k), (k-1, b-k+1)) = 1, \quad 1 \leq k \leq b, \ n \in \mathbb{N},$$

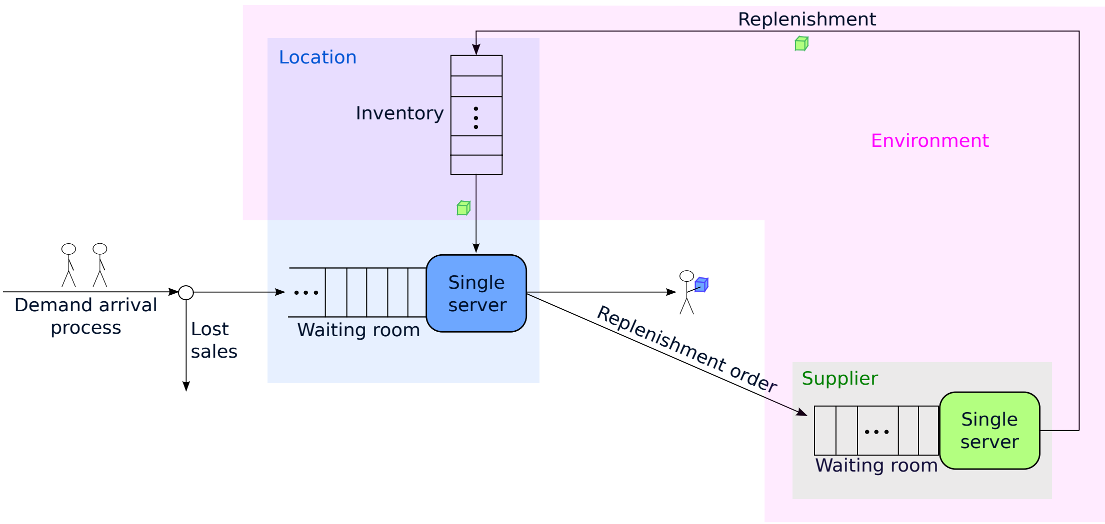


Figure D.1.2.: Production-inventory system with base stock policy

$$v_0((k,b-k),(\ell,b-\ell)) = \begin{cases} \nu, & \text{if } 0 \leq k < b,\ \ell = k+1, \\ \gamma \cdot k, & \text{if } 0 < k \leq b,\ \ell = k-1, \\ 0, & \text{otherwise for } k \neq \ell, \end{cases}$$

$$v_n((k,b-k),(\ell,b-\ell)) = \begin{cases} \nu, & \text{if } 0 \leq k < b,\ \ell = k+1, \\ \gamma \cdot (k-1), & \text{if } 0 < k \leq b,\ \ell = k-1, \\ 0, & \text{otherwise for } k \neq \ell, \qquad n \geq 1. \end{cases}$$

(b) The queueing-inventory system with phase-type distributed life times described in Section 4.2.2 fits into the definition of the queueing system in a random environment by setting

$$\begin{aligned} K &= \Big\{ \Big(\bar{k}, [h_1, \ldots, h_k], b - (\bar{k}+k)\Big) : \\ &\qquad \bar{k} \in \{0,1\}, h_\ell \in \{1,\ldots,H\},\ 1 \leq \ell \leq k,\ h_1 \leq h_2 \leq \cdots \leq h_k,\ 0 \leq \bar{k}+k \leq b \Big\}, \\ K_B &= \Big\{ \Big(\bar{k}, [h_1, \ldots, h_k], b - (\bar{k}+k)\Big) : \bar{k}+k = 0 \Big\} = \{(0,[0],b)\}, \\ K_W &= \Big\{ \Big(\bar{k}, [h_1, \ldots, h_k], b - (\bar{k}+k)\Big) : \bar{k}+k > 0 \Big\}, \end{aligned}$$

where K_W indicates for the inventory that there is stock on hand for production, and

$$R_n((0,[0],b),(0,[0],b)) = 1, \qquad R_n(\mathbf{k}; T_{0\searrow}\mathbf{k}) = 1, \quad \bar{k} = 1,\ n \in \mathbb{N},$$

for $n = 0$

$$v_0\Big(\Big(0,[0],b\Big);\Big(0,[\tilde{h}],b-1\Big)\Big) = \nu \cdot b(\tilde{h}),$$

for $n > 0$

$$v_n\Big(\big(0,[0],b\big);\big(1,[0],b-1\big)\Big) = \nu,$$

$$v_n\Big(\big(1,[0],b\big);\big(1,[\widetilde{h}],b-1\big)\Big) = \nu \cdot b(\widetilde{h}) \cdot 1_{\underbrace{\{0<\overline{k}+k<b\}}_{=\{b>1\}}},$$

for $n \geq 0$

$$v_n\Big(\mathbf{k};\widetilde{T}_{\widetilde{h}\to 1}\mathbf{k}\Big) = \nu \cdot b(\widetilde{h}) \cdot 1_{\{\widetilde{h}<h_1\}} \cdot 1_{\{0<\overline{k}+k<b\}} \cdot 1_{\{k>0\}},$$

$$v_n\Big(\mathbf{k};\widetilde{T}_{\widetilde{h}\to k+1}\mathbf{k}\Big) = \nu \cdot b(\widetilde{h}) \cdot 1_{\{h_k\leq\widetilde{h}\}} \cdot 1_{\{0<\overline{k}+k<b\}} \cdot 1_{\{k>0\}},$$

$$v_n\Big(\mathbf{k};\widetilde{T}_{\widetilde{h}\to \ell}\mathbf{k})\Big) = \nu \cdot b(\widetilde{h}) \cdot 1_{\{h_{\ell-1}\leq\widetilde{h}<h_\ell\}} \cdot 1_{\{0<\overline{k}+k<b\}} \cdot 1_{\{k>1\}} \cdot 1_{\{\ell\in\{2,\ldots,k\}\}},$$

$$v_n\Big(\mathbf{k};T_{\ell\searrow}\mathbf{k}\Big) = \beta \cdot 1_{\{h_\ell=1\}} \cdot 1_{\{k>0\}},$$

$$v_n\Big(\mathbf{k};T_{\ell\to 1}\mathbf{k}\Big) = \beta \cdot 1_{\{0<h_\ell-1<h_1\}} \cdot 1_{\{k>0\}},$$

$$v_n\Big(\mathbf{k};T_{\ell\to m}\mathbf{k}\Big) = \beta \cdot 1_{\{0<h_{m-1}\leq h_\ell-1<h_m\}} \cdot 1_{\{k>1\}} \cdot 1_{\{m\in\{2,\ldots,\ell\}\}}.$$

D.1.1. Ergodicity

The stochastic queueing-environment process Z is a homogeneous Markov process and irreducible on the state space E, and has an infinitesimal generator $\mathbf{Q} = (q(z;\tilde{z}) : z,\tilde{z} \in E)$ with the following transition rates for $(n,k) \in E$:

$$\begin{aligned} q\left((n,k);(n+1,k)\right) &= \lambda(n) \cdot 1_{\{k\in K_W\}}, \\ q\left((n,k);(n-1,\ell)\right) &= \mu(n) \cdot R_n(k,\ell) \cdot 1_{\{n>0\}} \cdot 1_{\{k\in K_W\}}, \\ q\left((n,k);(n,\ell)\right) &= v_n(k,\ell), \qquad\qquad k \neq \ell. \end{aligned}$$

Furthermore, $q(z;\tilde{z}) = 0$ for any other pair $z \neq \tilde{z}$, and

$$q\left(z;z\right) = -\sum_{\substack{\tilde{z}\in E,\\ z\neq\tilde{z}}} q\left(z;\tilde{z}\right) \qquad \forall z \in E.$$

We will show a necessary condition for ergodicity in Proposition D.1.4. Furthermore, a sufficient condition for positive recurrence is shown in Proposition D.1.8.

The proof for the necessary condition depends on the following result.

Proposition D.1.3. *If the queueing-environment process Z is recurrent, then any solution $\mathbf{x} = (x\left(z\right) : z \in E)$ of the global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ fulfils for all $n \in \mathbb{N}_0$*

$$\sum_{k\in K_W} x(n,k) = \sum_{k\in K_W} x(n+1,k) \cdot \frac{\mu(n+1)}{\lambda(n)} \tag{D.1.1}$$

and

$$\sum_{k\in K_W} x(n,k) = \sum_{k\in K_W} x(0,k)\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}. \tag{D.1.2}$$

Proof. From irreducibility and recurrence of Z follows that there exists one, and up to a multiplicative factor only one, stationary measure $\mathbf{x} = (x(z) : z \in E)$. This stationary measure $\mathbf{x}$ has the property $x(z) > 0$ for all $z \in E$ and can be found as a solution of the global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ (cf. [Asm03, Theorem 4.2, p. 51]).

The global balance equations $\mathbf{x}\cdot\mathbf{Q} = \mathbf{0}$ of the stochastic queueing-environment process Z are for $(n,k)\in E$ given by

$$\begin{aligned} & x(n,k)\Big(\lambda(n)\cdot 1_{\{k\in K_W\}} + \underbrace{\sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)}_{=-v_n(k,k)} + \mu(n)\cdot 1_{\{n>0\}}\cdot 1_{\{k\in K_W\}}\Big) \\ &= x(n-1,k)\cdot\lambda(n-1)\cdot 1_{\{n>0\}}\cdot 1_{\{k\in K_W\}} \\ &\quad + \sum_{\ell\in K_W} x(n+1,\ell)\cdot\mu(n+1)\cdot R_{n+1}(\ell,k) \\ &\quad + \sum_{\ell\in K\setminus\{k\}} x(n,\ell)\cdot v_n(\ell,k). \end{aligned} \tag{D.1.3}$$

Equation (D.1.1) can be proven by the cut-criterion for recurrent processes, which is presented in Theorem A.1.1(b). For $n\in\mathbb{N}_0$, equation (D.1.1) can be proven by a cut, which divides E into complementary sets according to the queue length of customers that is less than or equal to n or greater than n, i.e. into the sets

$$\Big\{(m,k) : m\in\{0,1,\dots,n\},\ k\in K\Big\},$$
$$\Big\{(\widetilde{m},\widetilde{k}) : \widetilde{m}\in\mathbb{N}_0\setminus\{0,1,\dots,n\},\ \widetilde{k}\in K\Big\},\quad n\in\mathbb{N}_0.$$

Then, the following holds for $n\in\mathbb{N}_0$

$$\begin{aligned} & \sum_{m=0}^{n}\sum_{k\in K}\sum_{\widetilde{m}=n+1}^{\infty}\sum_{\widetilde{k}\in K} x(m,k)\cdot q((m,k);(\widetilde{m},\widetilde{k})) \\ &= \sum_{\widetilde{m}=n+1}^{\infty}\sum_{\widetilde{k}\in K}\sum_{m=0}^{n}\sum_{k\in K} x(\widetilde{m},\widetilde{k})\cdot q((\widetilde{m},\widetilde{k});(m,k)) \\ \Leftrightarrow\quad & \sum_{k\in K_W} x(n,k)\cdot\lambda(n) = \sum_{\widetilde{k}\in K_W} x(n+1,\widetilde{k})\cdot\mu(n+1)\cdot\underbrace{\sum_{\ell\in K} R_{n+1}(\widetilde{k},\ell)}_{=1} \\ \Leftrightarrow\quad & \sum_{k\in K_W} x(n+1,k) = \sum_{k\in K_W} x(n,k)\cdot\frac{\lambda(n)}{\mu(n+1)}. \end{aligned}$$

Consequently, for $n\in\mathbb{N}_0$ follows

$$\sum_{k\in K_W} x(n,k) = \sum_{k\in K_W} x(0,k)\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)}.$$

□

Proposition D.1.4. *If the queueing-environment process Z is ergodic, it holds* $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} < \infty$.

Proof. If the queueing-environment process Z is ergodic, the normalisation constant, as the sum of the solution $\mathbf{x} = (x(z) : z \in E)$ of the global balance equations (D.1.3) for Z, is finite:

$$\sum_{n=0}^{\infty}\sum_{k\in K} x(n,k) < \infty.$$

It holds

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{k\in K} x(n,k)\\
=\ &\sum_{n=0}^{\infty}\sum_{k\in K_W} x(n,k) + \sum_{n=0}^{\infty}\sum_{k\in K_B} x(n,k)\\
\overset{(D.1.2)}{=}\ &\sum_{n=0}^{\infty}\underbrace{\sum_{k\in K_W} x(0,k)}_{=:\widetilde{W}}\cdot\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} + \sum_{n=0}^{\infty}\sum_{k\in K_B} x(n,k)\\
=\ &\widetilde{W}\cdot\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} + \sum_{n=0}^{\infty}\sum_{k\in K_B} x(n,k).
\end{aligned}$$

Because of ergodicity, $\widetilde{W} \in (0,\infty)$ and $\sum_{n=0}^{\infty}\sum_{k\in K_B} x(n,k) < \infty$.
Hence, $\sum_{n=0}^{\infty}\sum_{k\in K} x(n,k)$ is finite only if $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} < \infty$. □

A sufficient condition for positive recurrence of a Markov process can be shown by the Foster-Lyapunov stability criterion, which is presented in Theorem A.1.2 on page 260. This leads to the sufficient condition for ergodicity in Proposition D.1.8.

Foss and his coauthors [FST12] analyse the stability of two-component Markov chains by using Lyapunov functions under the additional assumption that one component forms a Markov chain itself. In our investigations neither the environment process nor the queueing process is a Markov process itself.

We start with three preparatory lemmas.

Lemma D.1.5. *For an $M/M/1/\infty$ queue with queue-length-dependent arrival intensities $\lambda(n) > 0$ and service intensities $\mu(n) > 0$ let there exists a Lyapunov function $\widetilde{\mathcal{L}} : \mathbb{N}_0 \to \mathbb{R}_0^+$ with finite exception set $\widetilde{F}$ and constant $\widetilde{\varepsilon} > 0$, which satisfies the Foster-Lyapunov stability criterion. Then, the following inequalities are satisfied:*

$$\lambda(0)\cdot\left(\widetilde{\mathcal{L}}(1) - \widetilde{\mathcal{L}}(0)\right) \le -\widetilde{\varepsilon}, \qquad \textit{if } 0 \notin \widetilde{F}, \tag{D.1.4}$$

$$\lambda(n)\cdot\left(\widetilde{\mathcal{L}}(n+1) - \widetilde{\mathcal{L}}(n)\right) + \mu(n)\cdot\left(\widetilde{\mathcal{L}}(n-1) - \widetilde{\mathcal{L}}(n)\right) \le -\varepsilon, \quad \textit{if } n \notin \widetilde{F},\ n > 0. \tag{D.1.5}$$

Proof. It holds $\left(\mathbf{Q}\cdot\widetilde{\mathcal{L}}\right)(n) \le -\widetilde{\varepsilon}$ for $n \notin \widetilde{F}$ and therefore, we have

$$\begin{aligned}&\lambda(0)\cdot\left(\widetilde{\mathcal{L}}(1)-\widetilde{\mathcal{L}}(0)\right)\\ &= \sum_{m\in\mathbb{N}_0\setminus\{0\}} q\,(0;m)\cdot\left(\widetilde{\mathcal{L}}(m)-\widetilde{\mathcal{L}}(0)\right) = \left(\mathbf{Q}\cdot\widetilde{\mathcal{L}}\right)(0) \le -\widetilde{\varepsilon}\end{aligned}$$

if $n = 0 \notin \widetilde{F}$, and

$$\begin{aligned}&\lambda(n)\cdot\left(\widetilde{\mathcal{L}}(n+1)-\widetilde{\mathcal{L}}(n)\right)+\mu(n)\cdot\left(\widetilde{\mathcal{L}}(n-1)-\widetilde{\mathcal{L}}(n)\right)\\ &= \sum_{m\in\mathbb{N}_0\setminus\{n\}} q\,(n;m)\cdot\left(\widetilde{\mathcal{L}}(m)-\widetilde{\mathcal{L}}(n)\right) = \left(\mathbf{Q}\cdot\widetilde{\mathcal{L}}\right)(n) \le -\varepsilon\end{aligned}$$

if $n \notin \widetilde{F}$, $n > 0$. □

Lemma D.1.6. *We define the function $\tau_n : K \to \mathbb{R}_0^+$ for $n \in \mathbb{N}_0$, where $\tau_n(k) > 0$ for $k \in K_B$ is the mean first entrance time in K_W when starting in $k \in K_B$ and for $\ell \in K_W$ we set $\tau_n(\ell) = 0$, since it means that entrance in K_W has already happened.*

For $k \in K_B$ holds

$$\tau_n(k) = \frac{1}{-v_n(k,k)} + \sum_{\ell\in K_B\setminus\{k\}} \frac{v_n(k,\ell)}{-v_n(k,k)}\cdot\tau_n(\ell),$$

which is equivalent to

$$-1 = \sum_{\ell\in K_B\setminus\{k\}} v_n(k,\ell)\cdot(\tau_n(\ell)-\tau_n(k)) - \sum_{\ell\in K_W} v_n(k,\ell)\cdot\tau_n(k), \qquad n\in\mathbb{N}_0. \qquad \text{(D.1.6)}$$

Proof. Because of the irreducibility of Z it holds the “flow condition” which is defined in [Kre16, Definition A.1.1, p. 195]:

$$\forall \widetilde{K}_B \subset K_B,\ \widetilde{K}_B \ne \emptyset: \quad \exists k \in \widetilde{K}_B,\ \ell \in \widetilde{K}_B^c: \quad v_n(k,\ell) > 0.$$

Consequently, if $\widetilde{K}_B = K_B$, then for all $n \in \mathbb{N}_0$ there exists some $k \in K_B$ and $\ell = \ell(n) \in K_W$ such that $v_n(k,\ell) > 0$ because of irreducibility.

Because of the finite state space $K := K_W \uplus K_B$ and the flow condition, for every starting point $k \in K_B$ the environment process will enter K_W after a positive finite mean time. Hence, for all $k \in K_B$ holds $0 < \tau_n(k) < \infty$.

The set $(\tau_n(k) : k \in K_B)$ satisfies the following set of first entrance equations:

$$\begin{aligned}
\tau_n(k) &= \sum_{\ell \in K_B \setminus \{k\}} \frac{v_n(k,\ell)}{-v_n(k,k)} \cdot \left(\frac{1}{-v_n(k,k)} + \tau_n(\ell) \right) \\
&\quad + \sum_{\ell \in K_W} \frac{v_n(k,\ell)}{-v_n(k,k)} \cdot \left(\frac{1}{-v_n(k,k)} + \tau_n(\ell) \right) \\
&= \sum_{\ell \in K \setminus \{k\}} \frac{v_n(k,\ell)}{-v_n(k,k)} \cdot \frac{1}{-v_n(k,k)} \\
&\quad + \sum_{\ell \in K_B \setminus \{k\}} \frac{v_n(k,\ell)}{-v_n(k,k)} \cdot \tau_n(\ell) + \sum_{\ell \in K_W} \frac{v_n(k,\ell)}{-v_n(k,k)} \cdot \underbrace{\tau_n(\ell)}_{=0} \\
&\overset{(\star)}{=} \frac{1}{-v_n(k,k)} + \sum_{\ell \in K_B \setminus \{k\}} \frac{v_n(k,\ell)}{-v_n(k,k)} \cdot \tau_n(\ell), \qquad n \in \mathbb{N}_0. \qquad \text{(D.1.7)}
\end{aligned}$$

$(\star)$ holds because $V_n = (v_n(k,\ell) : k,\ell \in K)$ is a conservative generator, i.e. for $n \in \mathbb{N}_0$ holds

$$\sum_{\ell \in K} v_n(k,\ell) = 0 \quad \Leftrightarrow \quad \sum_{\ell \in K \setminus \{k\}} v_n(k,\ell) = -v_n(k,k).$$

(D.1.7) is equivalent to

$$\begin{aligned}
-1 &= \left(\sum_{\ell \in K_B \setminus \{k\}} v_n(k,\ell) \cdot \tau_n(\ell) \right) + v_n(k,k) \cdot \tau_n(k) \\
&= \sum_{\ell \in K_B \setminus \{k\}} v_n(k,\ell) \cdot \tau_n(\ell) - \sum_{\ell \in K \setminus \{k\}} v_n(k,\ell) \cdot \tau_n(k) \\
&= \sum_{\ell \in K_B \setminus \{k\}} v_n(k,\ell) \cdot (\tau_n(\ell) - \tau_n(k)) - \sum_{\ell \in K_W} v_n(k,\ell) \cdot \tau_n(k), \qquad n \in \mathbb{N}_0.
\end{aligned}$$

□

Lemma D.1.7. *We define for $n \in \mathbb{N}_0$*

$$\widehat{c}_n = \left[\min \left\{ \frac{1}{\max_{k \in K_W} \left\{ \mu(n+1) \cdot \sum_{\ell \in K_B} R_{n+1}(k,\ell) \cdot \tau_n(\ell) \right\}}, \frac{1}{\max_{k \in K_W} \left\{ \sum_{\ell \in K_B} v_n(k,\ell) \cdot \tau_n(\ell) \right\}} \right\} \right], \qquad \text{(D.1.8)}$$

where $\tau_n(\cdot)$ is defined in Lemma D.1.6.

It holds $0 < \widehat{c}_n < \infty$ for $n \in \mathbb{N}_0$.

Proof. First, we will show that $\widehat{c}_n > 0$. It holds

$$\begin{aligned}
\widehat{c}_n > 0 \quad &\Leftrightarrow \quad \frac{1}{\max_{k\in K_W}\left\{\mu(n+1)\cdot\sum_{\ell\in K_B} R_{n+1}(k,\ell)\cdot\tau_n(\ell)\right\}} > 0 \\
&\qquad \wedge \quad \frac{1}{\max_{k\in K_W}\left\{\sum_{\ell\in K_B} v_n(k,\ell)\cdot\tau_n(\ell)\right\}} > 0 \\
&\Leftrightarrow \quad 0 < \mu(n+1) < \infty \quad \wedge \quad 0 < \tau_n(\ell) < \infty,\ \ell\in K_B \\
&\qquad \wedge \quad 0 \le R_{n+1}(k,\ell) < \infty, \quad k\in K_W,\ \ell\in K_B \\
&\qquad \wedge \quad 0 \le v_n(k,\ell) < \infty, \quad k\in K_W,\ \ell\in K_B.
\end{aligned}$$

The service intensity $\mu(n+1)$ is positive and finite. Moreover, in Lemma D.1.6 we have shown that the mean first entrance time $\tau_n(k)$ for $k \in K_B$ is positive and finite. In addition, the elements $R_{n+1}(k,\ell) \ge 0$ of the stochastic matrix and the off-diagonal elements $v_n(k,\ell) \ge 0$ of the generator matrix for $k \in K_W$ and $\ell \in K_B$ are non-negative and finite. Hence, it holds $\widehat{c}_n > 0$.

Second, we will show that $\widehat{c}_n < \infty$. It holds

$$\begin{aligned}
\widehat{c}_n < \infty \quad &\Leftrightarrow \quad \frac{1}{\max_{k\in K_W}\left\{\mu(n+1)\cdot\sum_{\ell\in K_B} R_{n+1}(k,\ell)\cdot\tau_n(\ell)\right\}} < \infty \\
&\qquad \vee \quad \frac{1}{\max_{k\in K_W}\left\{\sum_{\ell\in K_B} v_n(k,\ell)\cdot\tau_n(\ell)\right\}} < \infty \\
&\Leftrightarrow \quad \max_{k\in K_W}\left\{\mu(n+1)\cdot\sum_{\ell\in K_B} R_{n+1}(k,\ell)\cdot\tau_n(\ell)\right\} \neq 0 \\
&\qquad \vee \quad \max_{k\in K_W}\left\{\sum_{\ell\in K_B} v_n(k,\ell)\cdot\tau_n(\ell)\right\} \neq 0 \\
&\Leftrightarrow \quad \exists\ k\in K_W,\ \ell\in K_B: \quad R_{n+1}(k,\ell)\neq 0 \quad \vee \quad v_n(k,\ell)\neq 0.
\end{aligned}$$

The last equivalence holds because $\mu(n+1) > 0$ and $\tau_n(\ell) > 0$ for $\ell \in K_B$.

Now, we will prove that it exists $k \in K_W,\ \ell \in K_B$ with $R_{n+1}(k,\ell) \neq 0$ or $v_n(k,\ell) \neq 0$.

Because of irreducibility, all states communicate with each other. In particular, any state $(n,\ell) \in E$ with $\ell \in K_B$ must be reachable from some state $(\widetilde{n}, k) \in E$ with $k \in K_W$ in a finite number of steps. In Figure D.1.3, we illustrate the possible transitions into state $(n,\ell) \in E$ with $\ell \in K_B$.

If the environment is in state $\ell \in K_B$, no service is provided to customers in the queue and newly arriving customers are lost ("lost sales"). Hence, no transitions from the states $(n+1,\ell)$ or $(n-1,\ell)$ to state (n,ℓ) are possible.

Furthermore, a transition from state $(n-1,k)$ with $k \in K_W$ to state (n,ℓ) with $\ell \in K_B$ for all $n \in \mathbb{N}$ is not possible, since the changes of the environment's status are not coupled with the arrival process (cf. transition rates on page 328).

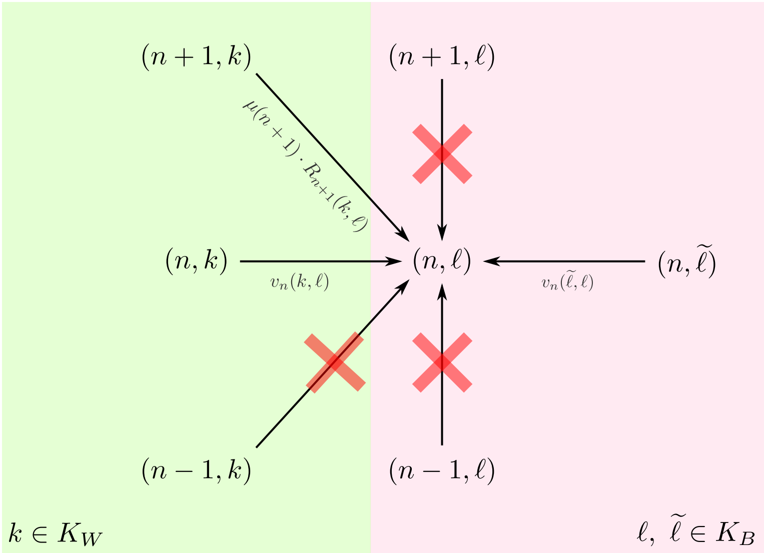


Figure D.1.3.: Typical transitions into state $(n,\ell) \in E$ with $\ell \in K_B$. Transitions marked by red crosses are impossible

The only remaining possible transitions are

- transitions from K_W to K_B with $v_n(k,\ell) > 0$ from some state $(n,k) \in \mathbb{N}_0 \times K_W$ to state $(n,\ell) \in \mathbb{N}_0 \times K_B$ or with $\mu(n+1) \cdot R_{n+1}(k,\ell) > 0$ from some state $(n+1,k) \in \mathbb{N}_0 \times K_W$ to state $(n,\ell) \in \mathbb{N}_0 \times K_B$ (see in Figure D.1.3)
- or transitions from inside of K_B with $v_n(\widetilde{\ell},\ell) > 0$ from some state $(n,\widetilde{\ell}) \in \mathbb{N}_0 \times K_B$ to state $(n,\ell) \in \mathbb{N}_0 \times K_B$.

Hence, as we mentioned before, because of irreducibility any state $(n,\ell) \in \mathbb{N}_0 \times K_B$ must be reachable from a state $(\widetilde{n},k) \in E$ with $\widetilde{n} \in \{n,n+1\}$ and $k \in K_W$ in a finite number of steps.

Definition. *A finite path from state $(n_0,k_0) \in E$ to state $(n_m,k_m) \in E$ is a sequence of vertices*

$$((n_0,k_0),\ (n_1,k_1),\ (n_2,k_2),\ldots,\ (n_{m-1},\ell_{m-1}),\ (n_m,\ell_m))\,,$$

where

$$q\left((n_i,k_i);(n_{i+1},k_{i+1})\right) > 0, \qquad i \in \{0,1,\ldots,m\}\,,$$

and m is the length of the path (number of edges).

Let $\mathfrak{M} \subseteq \mathbb{N}$ be the set of the lengths of the possible paths from states in $\{n, n+1\} \times K_W$ to state $(n_m, k_m) = (n, \ell) \in \mathbb{N}_0 \times K_B$. Let $m^* \geq 1$ be the minimal length.

Because of irreducibility and the abovementioned "only remaining possible transitions" there must exist a minimal path from a state in $\{n, n+1\} \times K_W$ to state $(n_m, k_m) = (n, \ell) \in \mathbb{N}_0 \times K_B$ of length $m^* \geq 1$ (see Figure D.1.4).

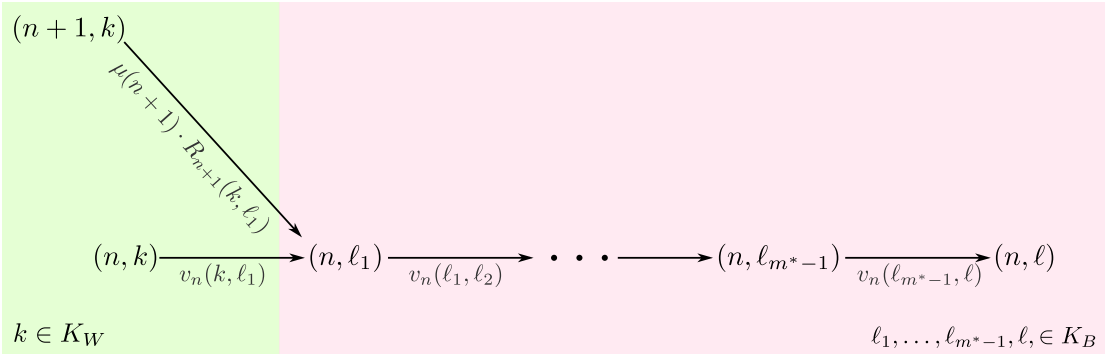


Figure D.1.4.: Path of minimal length m^*

If $m^* = 1$, the state $(n, \ell) \in \mathbb{N}_0 \times K_B$ is reachable directly from a state in $\{n, n+1\} \times K_W$ in one step.
Consequently, for $\ell \in K_B$ there exits $k \in K_W$ such that $R_{n+1}(k, \ell) > 0$ or $v_n(k, \ell) > 0$ holds, because these are the only possible transitions to reach the state $(n, \ell) \in \mathbb{N}_0 \times K_B$ from $\{n, n+1\} \times K_W$ in one step (see Figure D.1.3).

If $m^* > 1$, the state $(n, \ell) \in \mathbb{N}_0 \times K_B$ is reachable from a state in $\{n, n+1\} \times K_W$ in m^* steps.
Hence, there exists a minimal path of length m^*, where the state $(n, \ell_1) \in \mathbb{N}_0 \times K_B$ is reachable directly from a state in $\{n, n+1\} \times K_W$ in one step.
Consequently, for $\ell_1 \in K_B$ there exits $k \in K_W$ so that $R_{n+1}(k, \ell_1) > 0$ or $v_n(k, \ell_1) > 0$ holds.
Furthermore, it holds $v_n(\ell_1, \ell_2)$, $v_n(\ell_2, \ell_3)$, $\dots$, $v_n(\ell_{m-1}, \ell) > 0$, because these are the only possible transitions to reach the state $(n, \ell) \in \mathbb{N}_0 \times K_B$ from state $(n, \ell_1) \in \mathbb{N}_0 \times K_B$ in $m^* - 1$ steps.

□

Proposition D.1.8. *The queueing-environment process Z is ergodic if for an $M/M/1/\infty$ queue with queue-length-dependent arrival intensities $\lambda(n) > 0$ and service intensities $\mu(n) > 0$ there exists a Lyapunov function $\widetilde{\mathcal{L}} : \mathbb{N}_0 \to \mathbb{R}_0^+$ with finite exception set $\widetilde{F}$ and constant $\widetilde{\varepsilon} > 0$, which satisfies the Foster-Lyapunov stability criterion, and*

$$\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0,$$

where $\widehat{c}_n$ is defined in Lemma D.1.7.

Proof. The sufficient condition for positive recurrence can be shown by the Foster-Lyapunov stability criterion. We will show that $\mathcal{L} : E \to \mathbb{R}_0^+$ with

$$\mathcal{L}(n,k) = \widetilde{\mathcal{L}}(n) + 1_{\{k\in K_B\}} \cdot c_n \cdot \tau_n(k) \tag{D.1.9}$$

is a Lyapunov function with finite exception set $F = \widetilde{F} \times K$ and constant

$$\varepsilon = \min\left(\frac{\widetilde{\varepsilon}}{2}\,,\ \inf c_n\right) > 0,$$

where $c_n := \frac{\widetilde{\varepsilon}}{4} \cdot \widehat{c}_n$.

► First, we will check $(\mathbf{Q}\cdot\mathcal{L})(n,k) < \infty$ for $(n,k) \in F = \widetilde{F} \times K$.
Since $|K| < \infty$,
for $k \in K_B$ and $n \in \widetilde{F}$, $n \geq 0$, holds

$$\begin{aligned}(\mathbf{Q}\cdot\mathcal{L})(n,k) &= \sum_{(m,\ell)\in E\setminus\{(n,k)\}} q\left((n,k);(m,\ell)\right)\cdot\left(\mathcal{L}(m,\ell)-\mathcal{L}(n,k)\right)\\ &= \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot\left(\mathcal{L}(n,\ell)-\mathcal{L}(n,k)\right) < \infty,\end{aligned}$$

if $n = 0 \in \widetilde{F}$, then for $k \in K_W$ holds

$$\begin{aligned}(\mathbf{Q}\cdot\mathcal{L})(0,k) &= \sum_{(m,\ell)\in E\setminus\{(0,k)\}} q\left((0,k);(m,\ell)\right)\cdot\left(\mathcal{L}(m,\ell)-\mathcal{L}(0,k)\right)\\ &= \lambda(0)\cdot\left(\mathcal{L}(1,k)-\mathcal{L}(0,k)\right) + \sum_{\ell\in K\setminus\{k\}} v_0(k,\ell)\cdot\left(\mathcal{L}(0,\ell)-\mathcal{L}(0,k)\right) < \infty,\end{aligned}$$

for $k \in K_W$ and $n \in \widetilde{F}$, $n > 0$, holds

$$\begin{aligned}(\mathbf{Q}\cdot\mathcal{L})(n,k) &= \sum_{(m,\ell)\in E\setminus\{(n,k)\}} q\left((n,k);(m,\ell)\right)\cdot\left(\mathcal{L}(m,\ell)-\mathcal{L}(n,k)\right)\\ &= \lambda(n)\cdot\left(\mathcal{L}(n+1,k)-\mathcal{L}(n,k)\right)\\ &\quad + \sum_{\ell\in K}\mu(n)\cdot R_n(k,\ell)\cdot\left(\mathcal{L}(n-1,\ell)-\mathcal{L}(n,k)\right)\\ &\quad + \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot\left(\mathcal{L}(n,\ell)-\mathcal{L}(n,k)\right) < \infty.\end{aligned}$$

► Second, we will check $(\mathbf{Q}\cdot\mathcal{L})(n,k) \le -\varepsilon$ for $(n,k)\notin F = \widetilde{F}\times K$.
For $k\in K_B$ and $n\notin\widetilde{F}$, $n\ge 0$, holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})(n,k) = \sum_{(m,\ell)\in E\setminus\{(n,k)\}} q((n,k);(m,\ell))\cdot(\mathcal{L}(m,\ell)-\mathcal{L}(n,k))\\
&= \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot(\mathcal{L}(n,\ell)-\mathcal{L}(n,k))\\
&= \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot\Big(\Big(\widetilde{\mathcal{L}}(n)+1_{\{\ell\in K_B\}}\cdot c_n\cdot\tau_n(\ell)\Big)-\Big(\widetilde{\mathcal{L}}(n)+1_{\{k\in K_B\}}\cdot c_n\cdot\tau_n(k)\Big)\Big)\\
&= \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot 1_{\{\ell\in K_B\}}\cdot c_n\cdot\tau_n(\ell) - \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot\underbrace{1_{\{k\in K_B\}}}_{=1} c_n\cdot\tau_n(k)\\
&= \sum_{\ell\in K_B\setminus\{k\}} v_n(k,\ell)\cdot c_n\cdot\tau_n(\ell) - \sum_{\ell\in K\setminus\{k\}} v_n(k,\ell)\cdot c_n\cdot\tau_n(k)\\
&= \sum_{\ell\in K_B\setminus\{k\}} v_n(k,\ell)\cdot(c_n\cdot\tau_n(\ell)-c_n\cdot\tau_n(k)) - \sum_{\ell\in K_W} v_n(k,\ell)\cdot c_n\cdot\tau_n(k)\\
&= c_n\cdot\underbrace{\left[\sum_{\ell\in K_B\setminus\{k\}} v_n(k,\ell)\cdot(\tau_n(\ell)-\tau_n(k)) - \sum_{\ell\in K_W} v_n(k,\ell)\cdot\tau_n(k)\right]}_{\overset{(D.1.6)}{=}-1}\\
&= -c_n \le -\varepsilon = -\min\left(\frac{\widetilde{\varepsilon}}{2}\,,\ \inf c_n\right).
\end{aligned}
$$

If $n=0\notin\widetilde{F}$, then for $k\in K_W$ holds

$$
\begin{aligned}
&(\mathbf{Q}\cdot\mathcal{L})(n,k) = \sum_{(m,\ell)\in E\setminus\{(0,k)\}} q((0,k);(m,\ell))\cdot(\mathcal{L}(m,\ell)-\mathcal{L}(0,k))\\
&= \lambda(0)\cdot(\mathcal{L}(1,k)-\mathcal{L}(0,k))\\
&\quad + \sum_{\ell\in K\setminus\{k\}} v_0(k,\ell)\cdot(\mathcal{L}(0,\ell)-\mathcal{L}(0,k))\\
&= \lambda(0)\cdot\left(\left(\widetilde{\mathcal{L}}(1)+\underbrace{1_{\{k\in K_B\}}}_{=0}\cdot c_1\cdot\tau_1(k)\right)-\left(\widetilde{\mathcal{L}}(0)+\underbrace{1_{\{k\in K_B\}}}_{=0}\cdot c_0\cdot\tau_0(k)\right)\right)\\
&\quad + \sum_{\ell\in K\setminus\{k\}} v_0(k,\ell)\cdot\left(\left(\widetilde{\mathcal{L}}(0)+1_{\{\ell\in K_B\}}\cdot c_0\cdot\tau_0(\ell)\right)-\left(\widetilde{\mathcal{L}}(0)+\underbrace{1_{\{k\in K_B\}}}_{=0}\cdot c_0\cdot\tau_0(k)\right)\right)\\
&= \lambda(0)\cdot\left(\widetilde{\mathcal{L}}(1)-\widetilde{\mathcal{L}}(0)\right) + \sum_{\ell\in K\setminus\{k\}} v_0(k,\ell)\cdot\left(1_{\{\ell\in K_B\}}\cdot c_0\cdot\tau_0(\ell)\right)\\
&= \underbrace{\lambda(0)\cdot\left(\widetilde{\mathcal{L}}(1)-\widetilde{\mathcal{L}}(0)\right)}_{\overset{(\diamond)}{\le}-\widetilde{\varepsilon}} + \sum_{\ell\in K_B} v_0(k,\ell)\cdot c_0\cdot\tau_0(\ell)\\
&\le -\widetilde{\varepsilon} + \underbrace{\sum_{\ell\in K_B} v_0(k,\ell)\cdot c_0\cdot\tau_0(\ell)}_{\overset{(D.1.8)}{\le}\frac{\widetilde{\varepsilon}}{4}} \le -\frac{3}{4}\cdot\widetilde{\varepsilon} \le -\varepsilon = -\min\left(\frac{\widetilde{\varepsilon}}{2}\,,\ \inf c_n\right).
\end{aligned}
$$

($\diamondsuit$) holds because of (D.1.4) since $\widetilde{\mathcal{L}} : E \to \mathbb{R}_0^+$ is a Lyapunov function for the $M/M/1/\infty$ queue with queue-length-dependent arrival and service intensities with constant $\widetilde{\varepsilon}$.

For $k \in K_W$ and $n \notin \widetilde{F}$, $n > 0$, holds

$$
\begin{aligned}
&(\mathbf{Q} \cdot \mathcal{L})(n,k) = \sum_{(m,\ell)\in E\setminus\{(n,k)\}} q\left((n,k);(m,\ell)\right) \cdot (\mathcal{L}(m,\ell) - \mathcal{L}(n,k)) \\
&= \lambda(n) \cdot (\mathcal{L}(n+1,k) - \mathcal{L}(n,k)) \\
&\quad + \sum_{\ell \in K} \mu(n) \cdot R_n(k,\ell) \cdot (\mathcal{L}(n-1,\ell) - \mathcal{L}(n,k)) \\
&\quad + \sum_{\ell \in K\setminus\{k\}} v_n(k,\ell) \cdot (\mathcal{L}(n,\ell) - \mathcal{L}(n,k)) \\
&= \lambda(n) \cdot \left(\left(\widetilde{\mathcal{L}}(n+1) + \underbrace{1_{\{k\in K_B\}}}_{=0} \cdot c_{n+1} \cdot \tau_{n+1}(k) \right) - \left(\widetilde{\mathcal{L}}(n) + \underbrace{1_{\{k\in K_B\}}}_{=0} \cdot c_n \cdot \tau_n(k) \right) \right) \\
&\quad + \sum_{\ell \in K} \mu(n) \cdot R_n(k,\ell) \cdot \left(\left(\widetilde{\mathcal{L}}(n-1) + 1_{\{\ell\in K_B\}} \cdot c_{n-1} \cdot \tau_{n-1}(\ell) \right) \right. \\
&\qquad\qquad \left. - \left(\widetilde{\mathcal{L}}(n) + \underbrace{1_{\{k\in K_B\}}}_{=0} \cdot c_n \cdot \tau_n(k) \right) \right) \\
&\quad + \sum_{\ell \in K\setminus\{k\}} v_n(k,\ell) \cdot \left(\left(\widetilde{\mathcal{L}}(n) + 1_{\{\ell\in K_B\}} \cdot c_n \cdot \tau_n(\ell) \right) - \left(\widetilde{\mathcal{L}}(n) + \underbrace{1_{\{k\in K_B\}}}_{=0} \cdot c_n \cdot \tau_n(k) \right) \right) \\
&= \lambda(n) \cdot \left(\widetilde{\mathcal{L}}(n+1) - \widetilde{\mathcal{L}}(n) \right) + \mu(n) \cdot \underbrace{\sum_{\ell \in K} R_n(k,\ell)}_{=1} \cdot \left(\widetilde{\mathcal{L}}(n-1) - \widetilde{\mathcal{L}}(n) \right) \\
&\quad + \mu(n) \cdot \sum_{\ell \in K} R_n(k,\ell) \cdot 1_{\{\ell\in K_B\}} \cdot c_{n-1} \cdot \tau_{n-1}(\ell) + \sum_{\ell \in K\setminus\{k\}} v_n(k,\ell) \cdot \left(1_{\{\ell\in K_B\}} \cdot c_n \cdot \tau_n(\ell) \right) \\
&= \underbrace{\lambda(n) \cdot \left(\widetilde{\mathcal{L}}(n+1) - \widetilde{\mathcal{L}}(n) \right) + \mu(n) \cdot \left(\widetilde{\mathcal{L}}(n-1) - \widetilde{\mathcal{L}}(n) \right)}_{\overset{(\triangle)}{\leq} -\widetilde{\varepsilon}} \\
&\quad + \mu(n) \cdot \sum_{\ell \in K_B} R_n(k,\ell) \cdot c_{n-1} \cdot \tau_{n-1}(\ell) + \sum_{\ell \in K_B} v_n(k,\ell) \cdot c_n \cdot \tau_n(\ell) \\
&\leq -\widetilde{\varepsilon} + \underbrace{\mu(n) \cdot \sum_{\ell \in K_B} R_n(k,\ell) \cdot c_{n-1} \cdot \tau_{n-1}(\ell)}_{\overset{(D.1.8)}{\leq} \frac{\widetilde{\varepsilon}}{4}} + \underbrace{\sum_{\ell \in K_B} v_n(k,\ell) \cdot c_n \cdot \tau_n(\ell)}_{\overset{(D.1.8)}{\leq} \frac{\widetilde{\varepsilon}}{4}} \\
&\leq -\frac{\widetilde{\varepsilon}}{2} \leq -\varepsilon = -\min\left(\frac{\widetilde{\varepsilon}}{2},\ \inf c_n \right).
\end{aligned}
$$

($\triangle$) holds because of (D.1.5) since $\widetilde{\mathcal{L}} : E \to \mathbb{R}_0^+$ is a Lyapunov function for the $M/M/1/\infty$ queue with queue-length-dependent arrival and service intensities with constant $\widetilde{\varepsilon}$. □

Corollary D.1.9. *The queueing-environment process Z is ergodic if there exists $N \in \mathbb{N}_0$ such that* $\inf_{n \geq N} (\mu(n) - \lambda(n)) > 0$ *and*

$$\inf_{n \in \mathbb{N}_0} \widehat{c}_n > 0,$$

where $\widehat{c}_n$ is defined in Lemma D.1.7.

Proof. Let there be $N \in \mathbb{N}_0$ such that $\inf_{n \geq N} (\mu(n) - \lambda(n)) > 0$. Then, for every $M/M/1/\infty$ queue with queue-length-dependent arrival intensities $\lambda(n) > 0$ and service intensities $\mu(n) > 0$ is $\widetilde{\mathcal{L}} : \mathbb{N}_0 \to \mathbb{R}_0^+$ with $\widetilde{\mathcal{L}}(n) = n$, finite exception set $\widetilde{F} = \{0, \dots, N-1\}$ and constant $\widetilde{\varepsilon} = \inf_{n \geq N} (\mu(n) - \lambda(n)) > 0$ a Lyapunov function, which satisfies the Foster-Lyapunov stability criterion. Hence, we can apply Proposition D.1.8. □

Remark D.1.10. The queueing-environment system can be modelled as a level-dependent quasi-birth-and-death process (LDQBD process). Under the assumptions from the above proposition, the queueing-environment system is ergodic and hence, we can use the algorithm of Bright and Taylor [BT95] to calculate the equilibrium distributions in LDQBD processes.

Example D.1.11. If $\sup_{n \in \mathbb{N}_0} \mu(n) < \infty$, then in the following examples holds $\inf_{n \in \mathbb{N}_0} \widehat{c}_n > 0$. It should be noted, that $\widehat{c}_n$ is only defined by $\mu(n+1)$, the generator V_n and the stochastic matrix R_n (since τ_n is defined by the generator V_n).

(a) For the generator $V_n = (v_n(k, \ell) : k, \ell \in K)$ holds

$$V_{2n+1} = V_1 \quad \text{and} \quad V_{2n} = V_2, \qquad n \in \mathbb{N}_0,$$

and for the stochastic matrix $R_n = (R_n(k, \ell) : k, \ell \in K)$ holds

$$R_{2n-1} = R_1 \quad \text{and} \quad R_{2n} = R_2, \qquad n \in \mathbb{N}.$$

A similar structure is found in birth-and-death processes with alternating rates which are considered e.g. by Di Crescenzo et al. ([DCMM14], [DCIM12]).

(b) For the generator $V_n = (v_n(k, \ell) : k, \ell \in K)$ holds

$$V_n = V \quad \forall n \geq N_0,$$

and for the stochastic matrix $R_n = (R_n(k, \ell) : k, \ell \in K)$ holds

$$R_n = R \quad \forall n \geq N_0.$$

Then, $\widehat{c}_n$ can be arbitrarily for $n \leq N_0 - 1$ and from N_0 it is bounded below by a $\widehat{c}_{\min} > 0$.
A similar structure is found in multi-server models ($M/M/s$ queue) which are studied e.g. by Neuts [Neu81, Section 6.2, Section 6.5].
Moreover, such a structure is found in a queue with N servers subject to breakdowns and repairs. This is studied by Neuts and Lucantoni [NL79].
Furthermore, this structure can be found in the study of Neuts and Rao [NR90]. Since an analytical solution for their multi-server retrial model is difficult, they make a simplifying approximation, which yields an infinitesimal generator with a modified matrix-geometric steady state vector which can be efficiently computed.
Additionally, such a structure with $N_0 = 1$ can be found in the queueing-inventory systems described in Section 4.2.3 and Section 4.2.2 (cf. Example D.1.2).

Special case: Queue-length-independent arrival and service intensities
In this paragraph, we analyse the queueing-inventory system with state-independent service rates μ and arrival rates λ in view of a sufficient and necessary criterion for ergodicity.

Necessity follows from Proposition D.1.4 with $\sum_{n=0}^{\infty}\prod_{m=1}^{n}\frac{\lambda(m-1)}{\mu(m)} = \sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n < \infty$.

Proposition D.1.12. *If the queueing-environment process Z is ergodic, it holds $\lambda < \mu$.*

Sufficiency follows from Proposition D.1.8 and Corollary D.1.9.

Proposition D.1.13.

(a) *The queueing-environment process Z is ergodic if for an $M/M/1/\infty$ queue with arrival rate $\lambda > 0$ and service rate $\mu > 0$ there exists a Lyapunov function $\widetilde{\mathcal{L}} : \mathbb{N}_0 \to \mathbb{R}_0^+$ with finite exception set $\widetilde{F}$ and constant $\widetilde{\varepsilon} > 0$, which satisfies the Foster-Lyapunov stability criterion, and*

$$\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0.$$

(b) *The queueing-environment process Z is ergodic if $\lambda < \mu$ and $\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0$.*

Example D.1.14. We consider systems with state-independent rates λ and μ.

(a) We first discuss the structure of the constants $\widehat{c}_n$, where the occurring levels are shifted: R_{n+1} versus v_n, $n \geq 0$.
We construct an example where

$$\max_{k\in K_W}\left\{\mu\cdot\sum_{\ell\in K_B} R_1(k,\ell)\cdot\tau_1(\ell)\right\} = 0 \quad\wedge\quad \max_{k\in K_W}\left\{\sum_{\ell\in K_B} v_1(k,\ell)\cdot\tau_1(\ell)\right\} = 0,$$

while

$$\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0.$$

We consider a queueing-environment system with $K_W = \{0\}$, $K_B = \{1\}$, stochastic matrix $R_n = (R_n(k,\ell) : k,\ell \in K)$ and generator matrix $V_n = (v_n(k,\ell) : k,\ell \in K)$ with

$$R_1 = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right), \qquad R_n = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1/2 & 1/2 \\ 1 & 0 & 1 \end{array}\right), \quad n \geq 2,$$

$$V_0 = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & -\eta & \eta \\ 1 & \eta & -\eta \end{array}\right), \qquad V_n = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & \eta & -\eta \end{array}\right), \quad n \geq 1.$$

The queueing-environment process is irreducible because all states communicate with each other (see the state transition diagram in Figure D.1.5).

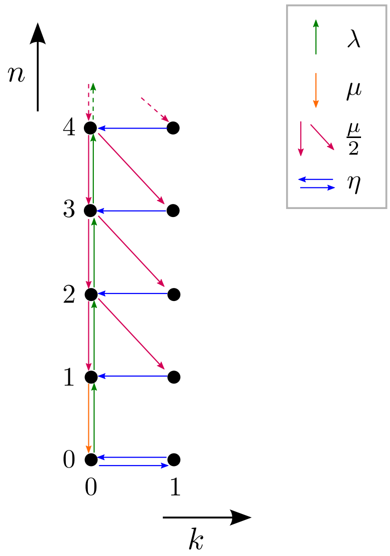


Figure D.1.5.: State transition diagram

For such a system it holds

$$\widehat{c}_0 = \min\left\{\frac{1}{\mu\cdot R_1(0,1)\cdot\tau_0(1)}, \frac{1}{v_0(0,1)\cdot\tau_0(1)}\right\} = \min\left\{\frac{1}{\mu\cdot 0\cdot\frac{1}{\eta}}, \frac{1}{\eta\cdot\frac{1}{\eta}}\right\}$$
$$= \min\{\infty, 1\} = 1,$$

$$\widehat{c}_1 = \min\left\{\frac{1}{\mu\cdot R_2(0,1)\cdot\tau_1(1)}, \frac{1}{v_1(0,1)\cdot\tau_1(1)}\right\} = \min\left\{\frac{1}{\mu\cdot\frac{1}{2}\cdot\frac{1}{\eta}}, \frac{1}{0\cdot\frac{1}{\eta}}\right\}$$
$$= \min\left\{\frac{2\cdot\eta}{\mu}, \infty\right\} = \frac{2\cdot\eta}{\mu},$$

$$\widehat{c}_n = \min\left\{\frac{1}{\mu\cdot R_{n+1}(0,1)\cdot\tau_n(1)}, \frac{1}{v_n(0,1)\cdot\tau_n(1)}\right\} = \min\left\{\frac{1}{\mu\cdot\frac{1}{2}\cdot\frac{1}{\eta}}, \frac{1}{0\cdot\frac{1}{\eta}}\right\} =$$
$$= \min\left\{\frac{2\cdot\eta}{\mu}, \infty\right\} = \frac{2\cdot\eta}{\mu}.$$

Hence,

$$\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0.$$

(b) We next discuss the condition $\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0$ in Proposition D.1.8.
We can modify example **(a)** such that $\inf_{n\in\mathbb{N}_0} \widehat{c}_n = 0$ while the queueing-environment process is irreducible.
We consider a queueing-environment system with $K_W = \{0\}$, $K_B = \{1\}$, stochastic matrix $\breve{R}_n = (\breve{R}_n(k,\ell) : k,\ell \in K)$ and generator matrix $\breve{V}_n = (\breve{v}_n(k,\ell) : k,\ell \in K)$ with

$$\breve{R}_1 = R_1 = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right), \qquad \breve{R}_n = R_n = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1/2 & 1/2 \\ 1 & 0 & 1 \end{array}\right), \quad n \geq 2,$$

$$\breve{V}_0 = V_0 = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & -\eta & \eta \\ 1 & \eta & -\eta \end{array}\right), \qquad \breve{V}_1 = V_1 = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & \eta & -\eta \end{array}\right),$$

$$\breve{V}_n = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & \frac{1}{n\cdot\eta} & \frac{1}{-n\cdot\eta} \end{array}\right), \quad n \geq 2.$$

The queueing-environment process is still irreducible.
For such a system it holds

$$\begin{aligned}\widehat{c}_0 &= \min\left\{\frac{1}{\mu\cdot R_1(0,1)\cdot\tau_0(1)}, \frac{1}{v_0(0,1)\cdot\tau_0(1)}\right\} = \min\left\{\frac{1}{\mu\cdot 0\cdot\frac{1}{\eta}}, \frac{1}{\eta\cdot\frac{1}{\eta}}\right\} \\ &= \min\{\infty, 1\} = 1,\end{aligned}$$

$$\begin{aligned}\widehat{c}_1 &= \min\left\{\frac{1}{\mu\cdot R_2(0,1)\cdot\tau_1(1)}, \frac{1}{v_1(0,1)\cdot\tau_1(1)}\right\} = \min\left\{\frac{1}{\mu\cdot\frac{1}{2}\cdot\frac{1}{\eta}}, \frac{1}{0\cdot\frac{1}{\eta}}\right\} \\ &= \min\left\{\frac{2\cdot\eta}{\mu}, \infty\right\} = \frac{2\cdot\eta}{\mu},\end{aligned}$$

$$\begin{aligned}\widehat{c}_n &= \min\left\{\frac{1}{\mu\cdot R_{n+1}(0,1)\cdot\tau_n(1)}, \frac{1}{v_n(0,1)\cdot\tau_n(1)}\right\} = \min\left\{\frac{1}{\mu\cdot\frac{1}{2}\cdot n\cdot\eta}, \frac{1}{0\cdot n\cdot\eta}\right\} = \\ &= \min\left\{\frac{2}{\mu\cdot n\cdot\eta}, \infty\right\} = \frac{2}{\mu\cdot n\cdot\eta}.\end{aligned}$$

Thus,

$$\inf_{n\in\mathbb{N}_0} \widehat{c}_n = 0.$$

The condition $\inf_{n\in\mathbb{N}_0} \widehat{c}_n > 0$ is required in Proposition D.1.8 to work with the specific Lyapunov function (D.1.9).

D.1.2. Properties of the stationary system

We assume throughout this section that the queueing-environment process Z is ergodic.

Definition D.1.15. For the queueing-environment system Z in a state space E, whose limiting distribution exists, we define

$$\pi := (\pi(n,k) : (n,k) \in E), \quad \pi(n,k) := \lim_{t\to\infty} P(Z(t) = (n,k))$$

and the appropriate marginal distributions

$$\xi := (\xi(n) : n \in \mathbb{N}_0), \quad \xi(n) := \lim_{t\to\infty} P(X(t) = n),$$

$$\theta := (\theta(k) : k \in K), \quad \theta(k) := \lim_{t\to\infty} P(Y(t) = k).$$

Let (X, Y) be a random variable which is distributed according to the queueing-environment process in equilibrium. Then, X resp. Y are random variables which are distributed according to the marginal steady state probability for the queue length resp. for the environment.

Proposition D.1.16. *The queueing-environment process Z fulfils for all $n \in \mathbb{N}_0$*

$$P(X = n, Y \in K_W) = P(X = n+1, Y \in K_W) \cdot \frac{\mu(n+1)}{\lambda(n)} \tag{D.1.10}$$

and

$$P(X = n, Y \in K_W) = P(X = 0, Y \in K_W) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \tag{D.1.11}$$

Hence, the probability that the environment is in K_W fulfils

$$P(Y \in K_W) = P(X = 0, Y \in K_W) \cdot \sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}.$$

Proof. The normalisation constant, as the sum of the solution $\mathbf{x} = (x(z) : z \in E)$ of the global balance equations (D.1.3) for Z, is finite because the queueing-environment process Z is ergodic. Then, with

$$\frac{x(n,k)}{\sum_{n\in\mathbb{N}_0}\sum_{k\in K} x(n,k)} = \pi(n,k) = P(X = n, Y = k)$$

follows in steady state for $n \in \mathbb{N}_0$ from (D.1.1)

$$\underbrace{\sum_{k\in K_W} \pi(n,k)}_{=P(X=n, Y\in K_W)} = \underbrace{\sum_{k\in K_W} \pi(n+1,k)}_{=P(X=n+1, Y\in K_W)} \cdot \frac{\mu(n+1)}{\lambda(n)}$$

and from (D.1.2)

$$\underbrace{\sum_{k\in K_W} \pi(n,k)}_{=P(X=n, Y\in K_W)} = \underbrace{\sum_{k\in K_W} \pi(0,k)}_{=P(X=0, Y\in K_W)} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}. \tag{D.1.12}$$

Summing up equation (D.1.11) over $n \in \mathbb{N}_0$ yields

$$P(Y \in K_W) = \sum_{n=0}^{\infty} P(X = n, Y \in K_W) \overset{\text{(D.1.11)}}{=} P(X = 0, Y \in K_W) \cdot \sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}.$$

□

Corollary D.1.17. *For the conditional distribution of the queue length process conditioned on $\{Y \in K_W\}$ holds for $n \in \mathbb{N}_0$*

$$P(X = n | Y \in K_W) = P(X = 0 | Y \in K_W) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}$$

with

$$P(X = 0 | Y \in K_W) = \left(\sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \right)^{-1}.$$

This shows that the conditional queue length process under the condition that the environment is in K_W has in equilibrium the same structure as a birth-and-death process with queue-length-dependent intensities.

Proof. $P(Y \in K_W) > 0$ because of ergodicity and equation (D.1.11) imply for $n \in \mathbb{N}_0$

$$\begin{aligned} P(X = n | Y \in K_W) = \frac{P(X = n, Y \in K_W)}{P(Y \in K_W)} &\overset{\text{(D.1.11)}}{=} \frac{P(X = 0, Y \in K_W)}{P(Y \in K_W)} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \\ &= P(X = 0 | Y \in K_W) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \end{aligned}$$

and the normalisation condition leads to

$$P(X = 0 | Y \in K_W) = \left(\sum_{n=0}^{\infty} \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \right)^{-1}.$$

An alternative direct proof is presented in the proof of Lemma 2.1.11 in [Kre16, p. 17]. □

Proposition D.1.18. *The limiting and stationary distribution of the queueing-environment process Z is in general not of product form.*

Proof. If the stationary distribution has a product form, it holds for any $n \in \mathbb{N}_0$

$$P(X = n, Y \in K_W) = P(X = n) \cdot P(Y \in K_W).$$

Then, it follows from Corollary D.1.17

$$\begin{aligned} P(X = n) &= P(X = 0 | Y \in K_W) \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)} \\ &= \left(\sum_{\tilde{n}=0}^{\infty} \prod_{m=1}^{\tilde{n}} \frac{\lambda(m-1)}{\mu(m)} \right)^{-1} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}, \qquad n \in \mathbb{N}_0. \end{aligned}$$

Consequently, if the stationary distribution has a product form

$$\pi(n,k) = \xi(n) \cdot \theta(k), \qquad n \in \mathbb{N}_0,\ k \in K,$$

then

$$\xi(n) = C^{-1} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}, \qquad n \in \mathbb{N}_0,$$

with normalisation constant

$$C = \sum_{\widetilde{n}=0}^{\infty} \prod_{m=1}^{\widetilde{n}} \frac{\lambda(m-1)}{\mu(m)}.$$

It has to be shown that this distribution satisfies the global balance equations

$$\begin{aligned} &\pi(n,k) \cdot \Big(\lambda(n) \cdot 1_{\{k \in K_W\}} - v_n(k,k) + \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}}\Big) \\ &= \pi(n-1,k) \cdot \lambda(n-1) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}} + \sum_{\ell \in K_W} \pi(n+1,\ell) \cdot \mu(n+1) \cdot R_{n+1}(\ell,k) \\ &+ \sum_{\ell \in K \setminus \{k\}} \pi(n,\ell) \cdot v_n(\ell,k). \end{aligned}$$

Substitution of $\pi(n,k) = \xi(n) \cdot \theta(k)$ into the global balance equations directly leads to

$$\begin{aligned} &\xi(n) \cdot \theta(k) \cdot \Big(\lambda(n) \cdot 1_{\{k \in K_W\}} - v_n(k,k) + \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}}\Big) \\ &= \xi(n-1) \cdot \theta(k) \cdot \lambda(n-1) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}} + \sum_{\ell \in K_W} \xi(n+1) \cdot \theta(\ell) \cdot \mu(n+1) \cdot R_{n+1}(\ell,k) \\ &+ \sum_{\ell \in K \setminus \{k\}} \xi(n) \cdot \theta(\ell) \cdot v_n(\ell,k). \end{aligned}$$

By substitution of $\xi(n) = C^{-1} \cdot \prod_{m=1}^{n} \frac{\lambda(m-1)}{\mu(m)}$ we obtain

$$\begin{aligned} &\xi(n) \cdot \theta(k) \cdot \Big(\lambda(n) \cdot 1_{\{k \in K_W\}} - v_n(k,k) + \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}}\Big) \\ &= \xi(n) \cdot \theta(k) \cdot \mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}} + \sum_{\ell \in K_W} \xi(n) \cdot \theta(\ell) \cdot \lambda(n) \cdot R_{n+1}(\ell,k) \\ &+ \sum_{\ell \in K \setminus \{k\}} \xi(n) \cdot \theta(\ell) \cdot v_n(\ell,k). \end{aligned}$$

Cancelling $\xi(n)$ and the sum with the terms $\mu(n) \cdot 1_{\{n>0\}} \cdot 1_{\{k \in K_W\}}$ on both sides of the equation leads to

$$\begin{aligned} &\theta(k) \cdot \Big(\lambda(n) \cdot 1_{\{k \in K_W\}} - v_n(k,k)\Big) \\ &= \sum_{\ell \in K_W} \theta(\ell) \cdot \lambda(n) \cdot R_{n+1}(\ell,k) + \sum_{\ell \in K \setminus \{k\}} \theta(\ell) \cdot v_n(\ell,k). \end{aligned} \tag{D.1.13}$$

However, this stands in contradiction to the product form assumption since $\theta(k)$ cannot be defined independently of n. ☐

Special case: Queue-length-independent arrival and service rates
In this paragraph, we analyse the queueing-environment system with state-independent service rates μ and arrival rates λ. Recall that the queueing-environment process Z is ergodic.

The following proposition is a special case of Proposition D.1.16.

Proposition D.1.19. *The queueing-environment process Z fulfils for all $n \in \mathbb{N}_0$*

$$P(X = n, Y \in K_W) = P(X = n + 1, Y \in K_W) \cdot \frac{\mu}{\lambda} \tag{D.1.14}$$

and

$$P(X = n, Y \in K_W) = P(X = 0, Y \in K_W) \cdot \left(\frac{\lambda}{\mu}\right)^n. \tag{D.1.15}$$

The following corollary is a special case of Corollary D.1.17.

Corollary D.1.20. *For the conditional distribution of the queue length process conditioned on $\{Y \in K_W\}$ holds for $n \in \mathbb{N}_0$*

$$P\left(X = n | Y \in K_W\right) = P\left(X = 0 | Y \in K_W\right) \cdot \left(\frac{\lambda}{\mu}\right)^n$$

with

$$P\left(X = 0 | Y \in K_W\right) = \left(1 - \frac{\lambda}{\mu}\right).$$

This shows that the conditional queue length process under the condition that the environment is in K_W has in equilibrium the same structure as a birth-and-death process with positive birth-rates λ and positive death-rates μ.

Proposition D.1.21. *For the queueing-environment process Z holds the following equilibrium of probability flows*

$$\underbrace{P(Y \in K_W) \cdot \lambda}_{\text{effective arrival rate}} = \underbrace{P(X > 0, Y \in K_W) \cdot \mu}_{\text{effective departure rate}}.$$

Hence, the probability that the environment is in K_W fulfils

$$P(Y \in K_W) = P(X > 0, Y \in K_W) \cdot \frac{\mu}{\lambda} \tag{D.1.16}$$

and

$$P(Y \in K_W) = P(X = 0, Y \in K_W) \cdot \frac{\mu}{\mu - \lambda}. \tag{D.1.17}$$

Proof. Summation of equation (D.1.14) over $n \in \mathbb{N}_0$ yields

$$\begin{aligned} P(Y \in K_W) &= \sum_{n=0}^{\infty} P(X = n, Y \in K_W) \overset{(D.1.14)}{=} \sum_{n=0}^{\infty} P(X = n + 1, Y \in K_W) \cdot \frac{\mu}{\lambda} \\ &= P(X > 0, Y \in K_W) \cdot \frac{\mu}{\lambda} \end{aligned}$$

and summation of equation (D.1.15) over $n \in \mathbb{N}_0$ yields

$$\begin{aligned} P(Y \in K_W) &= \sum_{n=0}^{\infty} P(X = n, Y \in K_W) \overset{\text{(D.1.15)}}{=} \sum_{n=0}^{\infty} P(X = 0, Y \in K_W) \cdot \left(\frac{\lambda}{\mu}\right)^n \\ &= P(X = 0, Y \in K_W) \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n = P(X = 0, Y \in K_W) \cdot \frac{\mu}{\mu - \lambda}. \end{aligned}$$

□

Proposition D.1.22. *The limiting and stationary distribution of Z is in general not of product form.*

Proof. The structure of the proof is similar to the proof of Proposition D.1.18 for the queueing-environment system with queue-length-dependent arrival and service rates. If the stationary distribution has a product form, it holds for any $n \in \mathbb{N}_0$

$$P\left(X = n, Y > 0\right) = P\left(X = n\right) \cdot P\left(Y > 0\right).$$

Then, it follows from Corollary D.1.20 for $n \in \mathbb{N}_0$

$$\begin{aligned} P(X = n) &= \frac{P\left(X = n\right) \cdot P\left(Y > 0\right)}{P\left(Y > 0\right)} = \frac{P\left(X = n, Y > 0\right)}{P\left(Y > 0\right)} = P\left(X = n | Y > 0\right) \\ &= P\left(X = 0 | Y > 0\right) \cdot \left(\frac{\lambda}{\mu}\right)^n = \left(1 - \frac{\lambda}{\mu}\right) \cdot \left(\frac{\lambda}{\mu}\right)^n. \end{aligned}$$

Consequently, if the stationary distribution has a product form

$$\pi\left(n, k\right) = \xi(n) \cdot \theta\left(k\right), \qquad n \in \mathbb{N}_0,\ k \in K,$$

then

$$\xi(n) = C^{-1} \cdot \left(\frac{\lambda}{\mu}\right)^n, \qquad n \in \mathbb{N}_0,$$

with normalisation constant $C^{-1} = \left(1 - \frac{\lambda}{\mu}\right)$.

By substitution of this stationary distribution into the global balance equations we also get an equation as in (D.1.13) (with λ instead of $\lambda(n)$) which is in contradiction to the product form assumption. □

	Queueing system in a random environment		**Queueing-inventory system**			
			exp		**PH**	
	queue-length-dependent?		**queue-length-dependent?**		**queue-length-dependent?**	
	yes	**no**	**yes**	**no**	**yes**	**no**
Ergodicity	Proposition D.1.3		Proposition 4.2.19		Proposition 4.2.4	
	Proposition D.1.4	Proposition D.1.12	Proposition 4.2.20	Theorem 4.2.24	Proposition 4.2.5	Theorem 4.2.9
	Proposition D.1.8	Proposition D.1.13	Proposition 4.2.21		Proposition 4.2.6	
	Corollary D.1.9					
Properties of the stationary system	Proposition D.1.16	Proposition D.1.19	Proposition 4.2.26	Proposition 4.2.30	Proposition 4.2.11	Proposition 4.2.14
	Corollary D.1.17	Corollary D.1.20	Corollary 4.2.27	Corollary 4.2.31	Corollary 4.2.12	Corollary 4.2.15
		Proposition D.1.21		Proposition 4.2.32		Proposition 4.2.16
			Proposition 4.2.28	Proposition 4.2.34		
	Proposition D.1.18	Proposition D.1.22	Proposition 4.2.29	Proposition 4.2.35	Remark 4.2.13	Remark 4.2.18

Table D.1.: Corresponding results in Appendix D.1 ,Section 4.2.3 and Section 4.2.2

D.2. Comparing throughputs with different ageing regimes

The following proposition is Proposition 4.3.11.

Proposition. *Consider three (exponential) ergodic production-inventory systems with the same arrival rate λ, service rate μ, replenishment rate ν, individual ageing rate γ for items in the inventory which are subject to ageing.*
The ageing regimes of the systems are different, which results in different Markovian state processes which we denote by Z^o under ageing regime $(\gamma\cdot k)\cdot 1_{\{m=0\}}+(\gamma\cdot(k-1)_+)\cdot 1_{\{m>0\}}$, Z^- under ageing regime $\gamma\cdot k$, and Z^+ under ageing regime $\gamma\cdot(k-1)_+$. Then the following holds.

(a) *If v_n^- is isotone, then for all $(m,k)\in E$ holds*

$$v_n^-(m,k)\le v_n^o(m,k)\quad \forall n\in\mathbb{N}, \tag{D.2.1}$$

and consequently $TH^-\le TH^o$.

(b) *If v_n^+ is isotone, then for all $(m,k)\in E$ holds*

$$v_n^o(m,k)\le v_n^+(m,k)\quad \forall n\in\mathbb{N}, \tag{D.2.2}$$

and consequently $TH^o\le TH^+$.

(c) *If v_n^o is isotone, then for all $(m,k)\in E$ holds*

$$v_n^-(m,k)\le v_n^o(m,k)\le v_n^+(m,k)\quad \forall n\in\mathbb{N}, \tag{D.2.3}$$

and consequently $TH^-\le TH^o\le TH^+$.

Proof. We proceed by induction over the number of jumps of the uniformization chains Z_u^-, Z_u^o, Z_u^+ and compare the respective cumulative rewards. By definition is in any case

$$v_1^-(m,k)=v_1^o(m,k)=v_1^+(m,k)=r(m,k)\quad \forall (m,k)\in E.$$

Assume that for some $n\ge 1$ holds

$$v_n^-(m,k)\le v_n^o(m,k)\le v_n^+(m,k),\quad \forall (m,k)\in E.$$

To perform the induction step we have to show

$$v_{n+1}^-(m,k)\le v_{n+1}^o(m,k)\le v_{n+1}^+(m,k),\quad \forall (m,k)\in E.$$

By $v_{n+1}^*=r+R^*\cdot v_n^*$ for $*\in\{o,-,+\}$ this reduces to

$$(R^-v_n^-)(m,k)\le (R^ov_n^o)(m,k)\le (R^+v_n^+)(m,k),\quad \forall (m,k)\in E.$$

For states $(m,0)$, $m\in\mathbb{N}_0$, we have for $*\in\{o,-,+\}$

$$(R^*v_n^*)(m,0)=\frac{\nu}{\alpha}\cdot v_n^*(m,1)+\frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v_n^*(m,0),$$

which proves the induction step in this case for **(a)**, **(b)**, **(c)**. The other cases need more detailed arguments. We first compute expressions $(R^*v_n^*)(m,k)$ and discuss then the comparison arguments.

For state $(0, b)$ we have

$$(R^- v_n^-)(0,b) = \frac{\lambda}{\alpha} \cdot v_n^-(1,b) + \frac{\gamma \cdot b}{\alpha} \cdot v_n^-(0,b-1) + \frac{\mu+\nu}{\alpha} \cdot v_n^-(0,b), \tag{D.2.4}$$

$$(R^o v_n^o)(0,b) = \frac{\lambda}{\alpha} \cdot v_n^o(1,b) + \frac{\gamma \cdot b}{\alpha} \cdot v_n^o(0,b-1) + \frac{\mu+\nu}{\alpha} \cdot v_n^o(0,b), \tag{D.2.5}$$

$$(R^+ v_n^+)(0,b) = \frac{\lambda}{\alpha} \cdot v_n^+(1,b) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^+(0,b-1) + \frac{\mu+\nu+\gamma}{\alpha} \cdot v_n^+(0,b). \tag{D.2.6}$$

For states $(0, k)$ with $k \in \{1, \ldots, b-1\}$ we have

$$\begin{aligned}(R^- v_n^-)(0,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^-(1,k) + \frac{\nu}{\alpha} \cdot v_n^-(0,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v_n^-(0,k-1) \\ & + \frac{\mu + \gamma \cdot (b-k)}{\alpha} \cdot v_n^-(0,k),\end{aligned} \tag{D.2.7}$$

$$\begin{aligned}(R^o v_n^o)(0,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^o(1,k) + \frac{\nu}{\alpha} \cdot v_n^o(0,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v_n^o(0,k-1) \\ & + \frac{\mu + \gamma \cdot (b-k)}{\alpha} \cdot v_n^o(0,k),\end{aligned} \tag{D.2.8}$$

$$\begin{aligned}(R^+ v_n^+)(0,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^+(1,k) + \frac{\nu}{\alpha} \cdot v_n^+(0,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^+(0,k-1) \\ & + \frac{\mu + \gamma \cdot (b-k+1)}{\alpha} \cdot v_n^+(0,k).\end{aligned} \tag{D.2.9}$$

For states (m, b) with $m \in \mathbb{N}$ we have

$$\begin{aligned}(R^- v_n^-)(m,b) = {} & \frac{\lambda}{\alpha} \cdot v_n^-(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,b-1) \\ & + \frac{\gamma \cdot b}{\alpha} \cdot v_n^-(m,b-1) + \frac{\nu}{\alpha} \cdot v_n^-(m,b),\end{aligned} \tag{D.2.10}$$

$$\begin{aligned}(R^o v_n^o)(m,b) = {} & \frac{\lambda}{\alpha} \cdot v_n^o(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^o(m-1,b-1) \\ & + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^o(m,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v_n^o(m,b),\end{aligned} \tag{D.2.11}$$

$$\begin{aligned}(R^+ v_n^+)(m,b) = {} & \frac{\lambda}{\alpha} \cdot v_n^+(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^+(m-1,b-1) \\ & + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^+(m,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v_n^+(m,b).\end{aligned} \tag{D.2.12}$$

For states (m,k) with $m \in \mathbb{N}$ and $k \in \{1,\ldots,b-1\}$ we have

$$(R^- v_n^-)(m,k) = \frac{\lambda}{\alpha} \cdot v_n^-(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^-(m,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v_n^-(m,k-1) + \frac{\gamma \cdot (b-k)}{\alpha} \cdot v_n^-(m,k), \tag{D.2.13}$$

$$(R^o v_n^o)(m,k) = \frac{\lambda}{\alpha} \cdot v_n^o(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^o(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^o(m,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^o(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v_n^o(m,k), \tag{D.2.14}$$

$$(R^+ v_n^+)(m,k) = \frac{\lambda}{\alpha} \cdot v_n^+(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^+(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^+(m,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^+(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v_n^+(m,k). \tag{D.2.15}$$

(a) Comparing (D.2.4) and (D.2.5) resp. (D.2.7) and (D.2.8) shows that for initial states $(0,k)$ for all $k \in \{1,\ldots,b\}$ the proposed inequality $v_{n+1}^-(0,k) \le v_{n+1}^o(0,k)$ holds.

We rewrite (D.2.10) and (D.2.11) as

$$(R^- v_n^-)(m,b) = \frac{\lambda}{\alpha} \cdot v_n^-(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^-(m,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v_n^-(m,b) + \left[\frac{\gamma}{\alpha} \cdot v_n^-(m,b-1) - \frac{\gamma}{\alpha} \cdot v_n^-(m,b)\right],$$

$$(R^o v_n^o)(m,b) = \frac{\lambda}{\alpha} \cdot v_n^o(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^o(m-1,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^o(m,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v_n^o(m,b)$$

and rewrite (D.2.13) and (D.2.14) as

$$(R^- v_n^-)(m,k) = \frac{\lambda}{\alpha} \cdot v_n^-(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^-(m,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^-(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v_n^-(m,k) + \left[\frac{\gamma}{\alpha} \cdot v_n^-(m,k-1) - \frac{\gamma}{\alpha} \cdot v_n^-(m,k)\right],$$

$$(R^o v_n^o)(m,k) = \frac{\lambda}{\alpha} \cdot v_n^o(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^o(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^o(m,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^o(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v_n^o(m,k).$$

If v_n^- is isotone, the differences in the blue squared brackets are non-positive.
This proves **(a)**.

(b) Comparing (D.2.11) and (D.2.12) resp. (D.2.14) and (D.2.15) shows that for initial states (m,k) with $m \geq 1$ and $k \in \{1,\ldots,b\}$ the proposed inequality $v^o_{n+1}(m,k) \leq v^+_{n+1}(m,k)$ holds.

We rewrite (D.2.5) and (D.2.6) as

$$(R^o v^o_n)(0,b) = \frac{\lambda}{\alpha} \cdot v^o_n(1,b) + \frac{\gamma \cdot b}{\alpha} \cdot v^o_n(0,b-1) + \frac{\mu+\nu}{\alpha} \cdot v^o_n(0,b),$$

$$\begin{aligned}(R^+ v^+_n)(0,b) = &\frac{\lambda}{\alpha} \cdot v^+_n(1,b) + \frac{\gamma \cdot b}{\alpha} \cdot v^+_n(0,b-1) + \frac{\mu+\nu}{\alpha} \cdot v^+_n(0,b) \\ &+ \left[\frac{\gamma}{\alpha} \cdot v^+_n(0,b) - \frac{\gamma}{\alpha} \cdot v^+_n(0,b-1)\right]\end{aligned}$$

and rewrite (D.2.8) and (D.2.9) as

$$\begin{aligned}(R^o v^o_n)(0,k) = &\frac{\lambda}{\alpha} \cdot v^o_n(1,k) + \frac{\nu}{\alpha} \cdot v^o_n(0,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v^o_n(0,k-1) \\ &+ \frac{\mu + \gamma \cdot (b-k)}{\alpha} \cdot v^o_n(0,k),\end{aligned}$$

$$\begin{aligned}(R^+ v^+_n)(0,k) = &\frac{\lambda}{\alpha} \cdot v^+_n(1,k) + \frac{\nu}{\alpha} \cdot v^+_n(0,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v^+_n(0,k-1) \\ &+ \frac{\mu + \gamma \cdot (b-k)}{\alpha} \cdot v^+_n(0,k) + \left[\frac{\gamma}{\alpha} \cdot v^+_n(0,k) - \frac{\gamma}{\alpha} \cdot v^+_n(0,k-1)\right].\end{aligned}$$

If v^+_n is isotone, the differences in the blue squared brackets are non-negative.
This proves **(b)**.

(c) To prove the two-sided bounds $v^-_{n+1}(m,k) \leq v^o_{n+1}(m,k) \leq v^+_{n+1}(m,k)$ we first check again first (D.2.4) and (D.2.5) resp. (D.2.7) and (D.2.8) and see that for initial states $(0,k)$ for all $k \in \{1,\ldots,b\}$ the proposed inequality $v^-_{n+1}(0,k) \leq v^o_{n+1}(0,k)$ holds.
We rewrite (D.2.5) and (D.2.6) as

$$\begin{aligned}(R^o v^o_n)(0,b) = &\frac{\lambda}{\alpha} \cdot v^o_n(1,b) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(0,b-1) + \frac{\mu+\nu+\gamma}{\alpha} \cdot v^o_n(0,b) \\ &+ \left[\frac{\gamma}{\alpha} \cdot v^o_n(0,b-1) - \frac{\gamma}{\alpha} \cdot v^o_n(0,b)\right],\end{aligned}$$

$$(R^+ v^+_n)(0,b) = \frac{\lambda}{\alpha} \cdot v^+_n(1,b) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^+_n(0,b-1) + \frac{\mu+\nu+\gamma}{\alpha} \cdot v^+_n(0,b)$$

and rewrite (D.2.8) and (D.2.9) as

$$
\begin{aligned}
(R^o v_n^o)(0,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^o(1,k) + \frac{\nu}{\alpha} \cdot v_n^o(0,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^o(0,k-1) \\
& + \frac{\mu + \gamma \cdot (b-k+1)}{\alpha} \cdot v_n^o(0,k) + \left[\frac{\gamma}{\alpha} \cdot v_n^o(0,k-1) - \frac{\gamma}{\alpha} \cdot v_n^o(0,k)\right],
\end{aligned}
$$

$$
\begin{aligned}
(R^+ v_n^+)(0,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^+(1,k) + \frac{\nu}{\alpha} \cdot v_n^+(0,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^+(0,k-1) \\
& + \frac{\mu + \gamma \cdot (b-k+1)}{\alpha} \cdot v_n^+(0,k).
\end{aligned}
$$

If v_n^o is isotone, the differences in blue squared brackets are non-positive, which proves this part of **(c)**.

We next check (D.2.11) and (D.2.12) resp. (D.2.14) and (D.2.15) and see that for initial states (m,k) with $m \geq 1$ and $k \in \{1, \ldots, b\}$ the proposed inequality $v_{n+1}^o(m,k) \leq v_{n+1}^+(m,k)$ holds.
We rewrite (D.2.10) and (D.2.11) as

$$
\begin{aligned}
(R^- v_n^-)(m,b) = {} & \frac{\lambda}{\alpha} \cdot v_n^-(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,b-1) \\
& + \frac{\gamma \cdot b}{\alpha} \cdot v_n^-(m,b-1) + \frac{\nu}{\alpha} \cdot v_n^-(m,b),
\end{aligned}
$$

$$
\begin{aligned}
(R^o v_n^o)(m,b) = {} & \frac{\lambda}{\alpha} \cdot v_n^o(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^o(m-1,b-1) \\
& + \frac{\gamma \cdot b}{\alpha} \cdot v_n^o(m,b-1) + \frac{\nu}{\alpha} \cdot v_n^o(m,b) + \left[\frac{\gamma}{\alpha} \cdot v_n^o(m,b) - \frac{\gamma}{\alpha} \cdot v_n^o(m,b-1)\right]
\end{aligned}
$$

and rewrite (D.2.13) and (D.2.14) as

$$
\begin{aligned}
(R^- v_n^-)(m,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^-(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^-(m,k+1) \\
& + \frac{\gamma \cdot k}{\alpha} \cdot v_n^-(m,k-1) + \frac{\gamma \cdot (b-k)}{\alpha} \cdot v_n^-(m,k),
\end{aligned}
$$

$$
\begin{aligned}
(R^o v_n^o)(m,k) = {} & \frac{\lambda}{\alpha} \cdot v_n^o(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^o(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^o(m,k+1) \\
& + \frac{\gamma \cdot k}{\alpha} \cdot v_n^o(m,k-1) + \frac{\gamma \cdot (b-k)}{\alpha} \cdot v_n^o(m,k) \\
& + \left[\frac{\gamma}{\alpha} \cdot v_n^o(m,k) - \frac{\gamma}{\alpha} \cdot v_n^o(m,k-1)\right].
\end{aligned}
$$

If v_n^o is isotone, the differences in blue squared brackets are non-negative.
This proves the remaining part of **(c)**. □

The following proposition is Proposition 4.3.12.

Proposition. *If in the production-inventory system with ageing regime $k \to \gamma \cdot k$ we have $\lambda \leq \gamma$, then the finite time cumulative rewards $v_n^-(m,k)$ are isotone with respect to the natural order.*

Proof. We show by induction isotonicity in both arguments and that the increase is bounded. For all $n \in \mathbb{N}$ holds

$$v_n^-(m,k) - v_n^-(m,k-1) \geq 0, \qquad \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0, \tag{D.2.16}$$
$$v_n^-(m+1,k) - v_n^-(m,k) \geq 0, \qquad \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0, \tag{D.2.17}$$
$$v_n^-(m+1,k) - v_n^-(m,k) \leq \alpha, \qquad \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0, \tag{D.2.18}$$
$$v_n^-(m,k) - v_n^-(m,k-1) \leq \alpha, \qquad \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0. \tag{D.2.19}$$

For $n = 1$ we have $v_1^-(m,k) = r(m,k) = \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}},\ \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0$, so (D.2.16)-(D.2.19) are trivially true.

Assume that (D.2.16)-(D.2.19) hold for some $n \in \mathbb{N}$. We shall verify these properties for v_{n+1}^-. In any case we exploit $v_{n+1}^- = r + R^- v_n^-$, $n \geq 1$.

► First we check (D.2.16).
For $m = 0$ and $k = 1$ holds

$$\begin{aligned}
&v_{n+1}^-(0,1) - v_{n+1}^-(0,0)\\
&= \left[\frac{\lambda}{\alpha} \cdot v_n^-(1,1) + \frac{\nu}{\alpha} \cdot v_n^-(0,2) + \frac{\gamma}{\alpha} \cdot v_n^-(0,0) + \frac{\mu + \gamma \cdot (b-1)}{\alpha} \cdot v_n^-(0,1)\right]\\
&\quad - \left[\frac{\nu}{\alpha} \cdot v_n^-(0,1) + \frac{\lambda + \mu + \gamma \cdot b}{\alpha} \cdot v_n^-(0,0)\right]\\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(1,1) - v_n^-(1,0))}_{\geq 0,\ \text{by (D.2.16)}} + \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(1,0) - v_n^-(0,0))}_{\geq 0,\ \text{by (D.2.17)}} + \frac{\nu}{\alpha} \cdot \underbrace{(v_n^-(0,2) - v_n^-(0,1))}_{\geq 0,\ \text{by (D.2.16)}}\\
&\quad + \frac{\mu + \gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v_n^-(0,1) - v_n^-(0,0))}_{\geq 0,\ \text{by (D.2.16)}} \geq 0.
\end{aligned}$$

For $m = 0$ and $k \in \{2,\dots,b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^-(0,k) - v_{n+1}^-(0,k-1)\\
&= \left[\frac{\lambda}{\alpha} \cdot v_n^-(1,k) + \frac{\nu}{\alpha} \cdot v_n^-(0,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v_n^-(0,k-1) + \frac{\mu + \gamma \cdot (b-k)}{\alpha} \cdot v_n^-(0,k)\right]\\
&\quad - \left[\frac{\lambda}{\alpha} \cdot v_n^-(1,k-1) + \frac{\nu}{\alpha} \cdot v_n^-(0,k) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^-(0,k-2)\right.\\
&\quad \left. + \frac{\mu + \gamma \cdot (b-k+1)}{\alpha} \cdot v_n^-(0,k-1)\right]
\end{aligned}$$

$$= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(1,k) - v_n^-(1,k-1))}_{\geq 0, \text{ by (D.2.16)}} + \frac{\nu}{\alpha} \cdot \underbrace{(v_n^-(0,k+1) - v_n^-(0,k))}_{\geq 0, \text{ by (D.2.16)}}$$
$$+ \frac{\gamma \cdot (k-1)}{\alpha} \cdot \underbrace{(v_n^-(0,k-1) - v_n^-(0,k-2))}_{\geq 0, \text{ by (D.2.16)}}$$
$$+ \frac{\mu + \gamma \cdot (b-k)}{\alpha} \cdot \underbrace{(v_n^-(0,k) - v_n^-(0,k-1))}_{\geq 0, \text{ by (D.2.16)}} \geq 0.$$

For $m = 0$ and $k = b$ holds

$$v_{n+1}^-(0,b) - v_{n+1}^-(0,b-1)$$
$$= \left[\frac{\lambda}{\alpha} \cdot v_n^-(1,b) + \frac{\gamma \cdot b}{\alpha} \cdot v_n^-(0,b-1) + \frac{\mu + \nu}{\alpha} \cdot v_n^-(0,b)\right]$$
$$- \left[\frac{\lambda}{\alpha} \cdot v_n^-(1,b-1) + \frac{\nu}{\alpha} \cdot v_n^-(0,b) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^-(0,b-2) + \frac{\mu + \gamma}{\alpha} \cdot v_n^-(0,b-1)\right]$$
$$= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(1,b) - v_n^-(1,b-1))}_{\geq 0, \text{ by (D.2.16)}} + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v_n^-(0,b-1) - v_n^-(0,b-2))}_{\geq 0, \text{ by (D.2.16)}}$$
$$+ \frac{\mu}{\alpha} \cdot \underbrace{(v_n^-(0,b) - v_n^-(0,b-1))}_{\geq 0, \text{ by (D.2.16)}} \geq 0.$$

For $m \geq 1$ and $k = 1$ holds

$$v_{n+1}^-(m,1) - v_{n+1}^-(m,0)$$
$$= \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^-(m+1,1) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,0) + \frac{\nu}{\alpha} \cdot v_n^-(m,2) + \frac{\gamma}{\alpha} \cdot v_n^-(m,0)\right.$$
$$\left. + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^-(m,1)\right]$$
$$- \left[0 + \frac{\nu}{\alpha} \cdot v_n^-(m,1) + \frac{\lambda + \mu + \gamma \cdot b}{\alpha} \cdot v_n^-(m,0)\right]$$
$$= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(m+1,1) - v_n^-(m,1))}_{\geq 0, \text{ by (D.2.17)}} + \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(m,1) - v_n^-(m,0))}_{\geq 0, \text{ by (D.2.16)}} + \frac{\nu}{\alpha} \cdot \underbrace{(v_n^-(m,2) - v_n^-(m,1))}_{\geq 0, \text{ by (D.2.16)}}$$
$$+ \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v_n^-(m,1) - v_n^-(m,0))}_{\geq 0, \text{ by (D.2.16)}} + \underbrace{\mu - \frac{\mu}{\alpha} \cdot \underbrace{(v_n^-(m,0) - v_n^-(m-1,0))}_{\in[0,\alpha], \text{ by (D.2.17), (D.2.18)}}}_{\geq 0} \geq 0.$$

For $m \geq 1$ and $k \in \{2, \ldots, b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^-(m,k) - v_{n+1}^-(m,k-1) \\
=& \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^-(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^-(m,k+1) \\
&+ \frac{\gamma \cdot k}{\alpha} \cdot v_n^-(m,k-1) + \frac{\gamma \cdot (b-k)}{\alpha} \cdot v_n^-(m,k)\Big] \\
&- \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^-(m+1,k-1) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,k-2) + \frac{\nu}{\alpha} \cdot v_n^-(m,k) \\
&+ \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^-(m,k-2) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v_n^-(m,k-1)\Big] \\
=& \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(m+1,k) - v_n^-(m+1,k-1))}_{\geq 0, \text{ by (D.2.16)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v_n^-(m-1,k-1) - v_n^-(m-1,k-2))}_{\geq 0, \text{ by (D.2.16)}} \\
&+ \frac{\nu}{\alpha} \cdot \underbrace{(v_n^-(m,k+1) - v_n^-(m,k))}_{\geq 0, \text{ by (D.2.16)}} + \frac{\gamma \cdot (k-1)}{\alpha} \cdot \underbrace{(v_n^-(m,k-1) - v_n^-(m,k-2))}_{\geq 0, \text{ by (D.2.16)}} \\
&+ \frac{\gamma \cdot (b-k)}{\alpha} \cdot \underbrace{(v_n^-(m,k) - v_n^-(m,k-1))}_{\geq 0, \text{ by (D.2.16)}} \geq 0.
\end{aligned}$$

For $m \geq 1$ and $k = b$ holds

$$\begin{aligned}
&v_{n+1}^-(m,b) - v_{n+1}^-(m,b-1) \\
=& \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^-(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,b-1) + \frac{\gamma \cdot b}{\alpha} \cdot v_n^-(m,b-1) + \frac{\nu}{\alpha} v_n^-(m,b)\Big] \\
&- \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^-(m+1,b-1) + \frac{\mu}{\alpha} \cdot v_n^-(m-1,b-2) + \frac{\nu}{\alpha} \cdot v_n^-(m,b) \\
&+ \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^-(m,b-2) + \frac{\gamma}{\alpha} \cdot v_n^-(m,b-1)\Big] \\
=& \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^-(m+1,b) - v_n^-(m+1,b-1))}_{\geq 0, \text{ by (D.2.16)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v_n^-(m-1,b-1) - v_n^-(m-1,b-2))}_{\geq 0, \text{ by (D.2.16)}} \\
&+ \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v_n^-(m,b-1) - v_n^-(m,b-2))}_{\geq 0, \text{ by (D.2.16)}} \geq 0.
\end{aligned}$$

► Second we check (D.2.17).
For $m = 0$ and $k = 0$ holds

$$\begin{aligned}
&v_{n+1}^-(1,0) - v_{n+1}^-(0,0)\\
&= \left[\frac{\nu}{\alpha}\cdot v_n^-(1,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v_n^-(1,0)\right] - \left[\frac{\nu}{\alpha}\cdot v_n^-(0,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v_n^-(0,0)\right]\\
&= \frac{\nu}{\alpha}\cdot\underbrace{(v_n^-(1,1) - v_n^-(0,1))}_{\geq 0,\quad \text{by (D.2.17)}} + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot\underbrace{(v_n^-(1,0) - v_n^-(0,0))}_{\geq 0,\quad \text{by (D.2.17)}} \geq 0.
\end{aligned}$$

For $m = 0$ and $k \in \{1,\ldots,b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^-(1,k) - v_{n+1}^-(0,k)\\
&= \Bigg[\mu + \frac{\lambda}{\alpha}\cdot v_n^-(2,k) + \frac{\mu}{\alpha}\cdot v_n^-(0,k-1) + \frac{\nu}{\alpha}\cdot v_n^-(1,k+1) + \frac{\gamma\cdot k}{\alpha}\cdot v_n^-(1,k-1)\\
&\quad + \frac{\gamma\cdot(b-k)}{\alpha}\cdot v_n^-(1,k)\Bigg]\\
&\quad - \left[0 + \frac{\lambda}{\alpha}\cdot v_n^-(1,k) + \frac{\nu}{\alpha}\cdot v_n^-(0,k+1) + \frac{\gamma\cdot k}{\alpha}\cdot v_n^-(0,k-1) + \frac{\mu+\gamma\cdot(b-k)}{\alpha}\cdot v_n^-(0,k)\right]\\
&= \frac{\lambda}{\alpha}\cdot\underbrace{(v_n^-(2,k) - v_n^-(1,k))}_{\geq 0,\quad \text{by (D.2.17)}} + \frac{\nu}{\alpha}\cdot\underbrace{(v_n^-(1,k+1) - v_n^-(0,k+1))}_{\geq 0,\quad \text{by (D.2.17)}}\\
&\quad + \frac{\gamma\cdot k}{\alpha}\cdot\underbrace{(v_n^-(1,k-1) - v_n^-(0,k-1))}_{\geq 0,\quad \text{by (D.2.17)}} + \frac{\gamma\cdot(b-k)}{\alpha}\cdot\underbrace{(v_n^-(1,k) - v_n^-(0,k))}_{\geq 0,\quad \text{by (D.2.17)}}\\
&\quad \underbrace{+\mu - \frac{\mu}{\alpha}\cdot\underbrace{(v_n^-(0,k) - v_n^-(0,k-1))}_{\in[0,\alpha],\quad \text{by (D.2.16), (D.2.19)}}}_{\geq 0} \geq 0.
\end{aligned}$$

For $m = 0$ and $k = b$ holds

$$\begin{aligned}
&v_{n+1}^-(1,b) - v_{n+1}^-(0,b)\\
&= \left[\mu + \frac{\lambda}{\alpha}\cdot v_n^-(2,b) + \frac{\mu}{\alpha}\cdot v_n^-(0,b-1) + \frac{\gamma\cdot b}{\alpha}\cdot v_n^-(1,b-1) + \frac{\nu}{\alpha}\cdot v_n^-(1,b)\right]\\
&\quad - \left[0 + \frac{\lambda}{\alpha}\cdot v_n^-(1,b) + \frac{\gamma\cdot b}{\alpha}\cdot v_n^-(0,b-1) + \frac{\mu+\nu}{\alpha}\cdot v_n^-(0,b)\right]\\
&= \frac{\lambda}{\alpha}\cdot\underbrace{(v_n^-(2,b) - v_n^-(1,b))}_{\geq 0,\quad \text{by (D.2.17)}} + \frac{\nu}{\alpha}\cdot\underbrace{(v_n^-(1,b) - v_n^-(0,b))}_{\geq 0,\quad \text{by (D.2.17)}}\\
&\quad + \frac{\gamma\cdot b}{\alpha}\cdot\underbrace{(v_n^-(1,b-1) - v_n^-(0,b-1))}_{\geq 0,\quad \text{by (D.2.17)}} \underbrace{+\mu - \frac{\mu}{\alpha}\cdot\underbrace{(v_n^-(0,b) - v_n^-(0,b-1))}_{\in[0,\alpha],\quad \text{by (D.2.16), (D.2.19)}}}_{\geq 0} \geq 0.
\end{aligned}$$

For $m \geq 1$ and $k = 0$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m+1,0) - v_{n+1}^{-}(m,0)\\
&= \left[\frac{\nu}{\alpha} \cdot v_n^{-}(m+1,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} \cdot v_n^{-}(m+1,0)\right]\\
&\quad - \left[\frac{\nu}{\alpha} \cdot v_n^{-}(m,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} \cdot v_n^{-}(m,0)\right]\\
&= \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,1) - v_n^{-}(m,1)\right)}_{\geq 0, \text{ by (D.2.17)}} + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,0) - v_n^{-}(m,0)\right)}_{\geq 0, \text{ by (D.2.17)}} \geq 0.
\end{aligned}$$

For $m \geq 1$ and $k \in \{1,\dots,b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m+1,k) - v_{n+1}^{-}(m,k)\\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+2,k) + \frac{\mu}{\alpha} \cdot v_n^{-}(m,k-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m+1,k+1)\\
&\quad + \frac{\gamma\cdot k}{\alpha} \cdot v_n^{-}(m+1,k-1) + \frac{\gamma\cdot(b-k)}{\alpha} \cdot v_n^{-}(m+1,k)\\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,k+1)\\
&\quad + \frac{\gamma\cdot k}{\alpha} \cdot v_n^{-}(m,k-1) + \frac{\gamma\cdot(b-k)}{\alpha} \cdot v_n^{-}(m,k)\Big]\\
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+2,k) - v_n^{-}(m+1,k)\right)}_{\geq 0, \text{ by (D.2.17)}} + \frac{\mu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m,k-1) - v_n^{-}(m-1,k-1)\right)}_{\geq 0, \text{ by (D.2.17)}}\\
&\quad + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,k+1) - v_n^{-}(m,k+1)\right)}_{\geq 0, \text{ by (D.2.17)}} + \frac{\gamma\cdot k}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,k-1) - v_n^{-}(m,k-1)\right)}_{\geq 0, \text{ by (D.2.17)}}\\
&\quad + \frac{\gamma\cdot(b-k)}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,k) - v_n^{-}(m,k)\right)}_{\geq 0, \text{ by (D.2.17)}} \geq 0.
\end{aligned}$$

For $m \geq 1$ and $k = b$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m+1,b) - v_{n+1}^{-}(m,b)\\
&= \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+2,b) + \frac{\mu}{\alpha} \cdot v_n^{-}(m,b-1) + \frac{\gamma\cdot b}{\alpha} \cdot v_n^{-}(m+1,b-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m+1,b)\right]\\
&\quad - \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,b-1) + \frac{\gamma\cdot b}{\alpha} \cdot v_n^{-}(m,b-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,b)\right]\\
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+2,b) - v_n^{-}(m+1,b)\right)}_{\geq 0, \text{ by (D.2.17)}} + \underbrace{\frac{\mu}{\alpha} \cdot \left(v_n^{-}(m,b-1) - v_n^{-}(m-1,b-1\right)}_{\geq 0, \text{ by (D.2.17)}}\\
&\quad + \frac{\gamma\cdot b}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,b-1) - v_n^{-}(m,b-1)\right)}_{\geq 0, \text{ by (D.2.17)}} + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,b) - v_n^{-}(m,b)\right)}_{\geq 0, \text{ by (D.2.17)}} \geq 0.
\end{aligned}$$

► Third we check (D.2.18).
For $m = 0$ and $k = 0$ holds

$$\begin{aligned}
&v_{n+1}^-(1,0) - v_{n+1}^-(0,0)\\
&= \left[\frac{\nu}{\alpha}\cdot v_n^-(1,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v_n^-(1,0)\right] - \left[\frac{\nu}{\alpha}\cdot v_n^-(0,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v_n^-(0,0)\right]\\
&= \frac{\nu}{\alpha}\cdot\underbrace{(v_n^-(1,1) - v_n^-(0,1))}_{\leq\alpha,\ \text{by (D.2.18)}} + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot\underbrace{(v_n^-(1,0) - v_n^-(0,0))}_{\leq\alpha,\ \text{by (D.2.18)}} \leq \alpha.
\end{aligned}$$

For $m = 0$ and $k \in \{1,\dots,b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^-(1,k) - v_{n+1}^-(0,k)\\
&= \Bigg[\mu + \frac{\lambda}{\alpha}\cdot v_n^-(2,k) + \frac{\mu}{\alpha}\cdot v_n^-(0,k-1) + \frac{\nu}{\alpha}\cdot v_n^-(1,k+1) + \frac{\gamma\cdot k}{\alpha}\cdot v_n^-(1,k-1)\\
&\quad + \frac{\gamma\cdot(b-k)}{\alpha}\cdot v_n^-(1,k)\Bigg]\\
&\quad - \left[0 + \frac{\lambda}{\alpha}\cdot v_n^-(1,k) + \frac{\nu}{\alpha}\cdot v_n^-(0,k+1) + \frac{\gamma\cdot k}{\alpha}\cdot v_n^-(0,k-1) + \frac{\mu+\gamma\cdot(b-k)}{\alpha}\cdot v_n^-(0,k)\right]\\
&= \frac{\lambda}{\alpha}\cdot\underbrace{(v_n^-(2,k) - v_n^-(1,k))}_{\leq\alpha,\ \text{by (D.2.18)}} + \frac{\nu}{\alpha}\cdot\underbrace{(v_n^-(1,k+1) - v_n^-(0,k+1))}_{\leq\alpha,\ \text{by (D.2.18)}}\\
&\quad + \frac{\gamma\cdot k}{\alpha}\cdot\underbrace{(v_n^-(1,k-1) - v_n^-(0,k-1))}_{\leq\alpha,\ \text{by (D.2.18)}} + \frac{\gamma\cdot(b-k)}{\alpha}\cdot\underbrace{(v_n^-(1,k) - v_n^-(0,k))}_{\leq\alpha,\ \text{by (D.2.18)}}\\
&\quad \underbrace{+\,\mu - \frac{\mu}{\alpha}\cdot\underbrace{(v_n^-(0,k) - v_n^-(0,k-1))}_{\in[0,\alpha],\ \text{by (D.2.16), (D.2.19)}}}_{\leq\mu} \leq \alpha.
\end{aligned}$$

For $m = 0$ and $k = b$ holds

$$\begin{aligned}
&v_{n+1}^-(1,b) - v_{n+1}^-(0,b)\\
&= \left[\mu + \frac{\lambda}{\alpha}\cdot v_n^-(2,b) + \frac{\mu}{\alpha}\cdot v_n^-(0,b-1) + \frac{\gamma\cdot b}{\alpha}\cdot v_n^-(1,b-1) + \frac{\nu}{\alpha}\cdot v_n^-(1,b)\right]\\
&\quad - \left[0 + \frac{\lambda}{\alpha}\cdot v_n^-(1,b) + \frac{\gamma\cdot b}{\alpha}\cdot v_n^-(0,b-1) + \frac{\mu+\nu}{\alpha}\cdot v_n^-(0,b)\right]\\
&= \frac{\lambda}{\alpha}\cdot\underbrace{(v_n^-(2,b) - v_n^-(1,b))}_{\leq\alpha,\ \text{by (D.2.18)}} + \frac{\nu}{\alpha}\cdot\underbrace{(v_n^-(1,b) - v_n^-(0,b))}_{\leq\alpha,\ \text{by (D.2.18)}}\\
&\quad + \frac{\gamma\cdot b}{\alpha}\cdot\underbrace{(v_n^-(1,b-1) - v_n^-(0,b-1))}_{\leq\alpha,\ \text{by (D.2.18)}} + \underbrace{\mu - \frac{\mu}{\alpha}\cdot\underbrace{(v_n^-(0,b) - v_n^-(0,b-1))}_{\in[0,\alpha],\ \text{by (D.2.16), (D.2.19)}}}_{\leq\mu} \leq \alpha.
\end{aligned}$$

For $m \geq 1$ and $k = 0$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m+1,0) - v_{n+1}^{-}(m,0) \\
&= \left[\frac{\nu}{\alpha} \cdot v_n^{-}(m+1,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} \cdot v_n^{-}(m+1,0)\right] \\
&\quad - \left[\frac{\nu}{\alpha} \cdot v_n^{-}(m,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} \cdot v_n^{-}(m,0)\right] \\
&= \frac{\nu}{\alpha} v \underbrace{\left(v_n^{-}(m+1,1) - v_n^{-}(m,1)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,0) - v_n^{-}(m,0)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} \leq \alpha.
\end{aligned}$$

For $m \geq 1$ and $k \in \{1, \ldots, b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m+1,k) - v_{n+1}^{-}(m,k) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+2,k) + \frac{\mu}{\alpha} \cdot v_n^{-}(m,k-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m+1,k+1) \\
&\quad + \frac{\gamma\cdot k}{\alpha} \cdot v_n^{-}(m+1,k-1) + \frac{\gamma\cdot(b-k)}{\alpha} \cdot v_n^{-}(m+1,k)\Big] \\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,k+1) \\
&\quad + \frac{\gamma\cdot k}{\alpha} \cdot v_n^{-}(m,k-1) + \frac{\gamma\cdot(b-k)}{\alpha} \cdot v_n^{-}(m,k)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+2,k) - v_n^{-}(m+1,k)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} + \frac{\mu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m,k-1) - v_n^{-}(m-1,k-1)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} \\
&\quad + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,k+1) - v_n^{-}(m,k+1)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} + \frac{\gamma\cdot k}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,k-1) - v_n^{-}(m,k-1)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} \\
&\quad + \frac{\gamma\cdot(b-k)}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,k) - v_n^{-}(m,k)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} \leq \alpha.
\end{aligned}$$

For $m \geq 1$ and $k = b$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m+1,b) - v_{n+1}^{-}(m,b) \\
&= \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+2,b) + \frac{\mu}{\alpha} \cdot v_n^{-}(m,b-1) + \frac{\gamma\cdot b}{\alpha} \cdot v_n^{-}(m+1,b-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m+1,b)\right] \\
&\quad - \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,b-1) + \frac{\gamma\cdot b}{\alpha} \cdot v_n^{-}(m,b-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,b)\right] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+2,b) - v_n^{-}(m+1,b)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} + \frac{\mu}{\alpha} \cdot \left(v_n^{-}(m,b-1) - v_n^{-}(m-1,b-1)\right) \\
&\quad + \frac{\gamma\cdot b}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,b-1) - v_n^{-}(m,b-1)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^{-}(m+1,b) - v_n^{-}(m,b)\right)}_{\leq\alpha, \quad \text{by (D.2.18)}} \leq \alpha.
\end{aligned}$$

► Fourth we check (D.2.19).
For $m = 0$ and $k = 1$ holds

$$\begin{aligned}
&v^-_{n+1}(0,1) - v^-_{n+1}(0,0)\\
&= \left[\frac{\lambda}{\alpha}\cdot v^-_n(1,1) + \frac{\nu}{\alpha}\cdot v^-_n(0,2) + \frac{\gamma}{\alpha}\cdot v^-_n(0,0) + \frac{\mu+\gamma\cdot(b-1)}{\alpha}\cdot v^-_n(0,1)\right]\\
&\quad - \left[\frac{\nu}{\alpha}\cdot v^-_n(0,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha} v^-_n(0,0)\right]\\
&= \frac{\lambda}{\alpha}\cdot \underbrace{(v^-_n(1,1) - v^-_n(0,1))}_{\leq\alpha,\ \text{ by (D.2.18)}} + \frac{\lambda}{\alpha}\cdot \underbrace{(v^-_n(0,1) - v^-_n(0,0))}_{\leq\alpha,\ \text{ by (D.2.19)}} + \frac{\nu}{\alpha}\cdot \underbrace{(v^-_n(0,2) - v^-_n(0,1))}_{\leq\alpha,\ \text{ by (D.2.19)}}\\
&\quad + \frac{\mu+\gamma\cdot(b-1)}{\alpha}\cdot \underbrace{(v^-_n(0,1) - v^-_n(0,0))}_{\leq\alpha,\ \text{ by (D.2.19)}} \leq (2\lambda+\mu+\nu+\gamma(b-1)) \overset{(\lambda\leq\gamma)}{\leq} \alpha.
\end{aligned}$$

For $m = 0$ and $k \in \{2, \ldots, b-1\}$ holds

$$\begin{aligned}
&v^-_{n+1}(0,k) - v^-_{n+1}(0,k-1)\\
&= \left[\frac{\lambda}{\alpha}\cdot v^-_n(1,k) + \frac{\nu}{\alpha}\cdot v^-_n(0,k+1) + \frac{\gamma\cdot k}{\alpha}\cdot v^-_n(0,k-1) + \frac{\mu+\gamma\cdot(b-k)}{\alpha}\cdot v^-_n(0,k)\right]\\
&\quad - \left[\frac{\lambda}{\alpha}\cdot v^-_n(1,k-1) + \frac{\nu}{\alpha}\cdot v^-_n(0,k) + \frac{\gamma\cdot(k-1)}{\alpha}\cdot v^-_n(0,k-2)\right.\\
&\quad \left. + \frac{\mu+\gamma\cdot(b-k+1)}{\alpha}\cdot v^-_n(0,k-1)\right]\\
&= \frac{\lambda}{\alpha}\cdot \underbrace{(v^-_n(1,k) - v^-_n(1,k-1))}_{\leq\alpha,\ \text{ by (D.2.19)}} + \frac{\nu}{\alpha}\cdot \underbrace{(v^-_n(0,k+1) - v^-_n(0,k))}_{\leq\alpha,\ \text{ by (D.2.19)}}\\
&\quad + \frac{\gamma\cdot(k-1)}{\alpha}\cdot \underbrace{(v^-_n(0,k-1) - v^-_n(0,k-2))}_{\leq\alpha,\ \text{ by (D.2.19)}} + \frac{\mu+\gamma\cdot(b-k)}{\alpha}\cdot \underbrace{(v^-_n(0,k) - v^-_n(0,k-1))}_{\leq\alpha,\ \text{ by (D.2.19)}}\\
&\leq \alpha - \gamma.
\end{aligned}$$

For $m = 0$ and $k = b$ holds

$$\begin{aligned}
&v^-_{n+1}(0,b) - v^-_{n+1}(0,b-1)\\
&= \left[\frac{\lambda}{\alpha}\cdot v^-_n(1,b) + \frac{\gamma\cdot b}{\alpha}\cdot v^-_n(0,b-1) + \frac{\mu+\nu}{\alpha}\cdot v^-_n(0,b)\right]\\
&\quad - \left[\frac{\lambda}{\alpha}\cdot v^-_n(1,b-1) + \frac{\nu}{\alpha}\cdot v^-_n(0,b) + \frac{\gamma\cdot(b-1)}{\alpha}\cdot v^-_n(0,b-2) + \frac{\mu+\gamma}{\alpha}\cdot v^-_n(0,b-1)\right]\\
&= \frac{\lambda}{\alpha}\cdot \underbrace{(v^-_n(1,b) - v^-_n(1,b-1))}_{\leq\alpha,\ \text{ by (D.2.19)}} + \frac{\gamma\cdot(b-1)}{\alpha}\cdot \underbrace{(v^-_n(0,b-1) - v^-_n(0,b-2))}_{\leq\alpha,\ \text{ by (D.2.19)}}\\
&\quad + \frac{\mu}{\alpha}\cdot \underbrace{(v^-_n(0,b) - v^-_n(0,b-1))}_{\leq\alpha,\ \text{ by (D.2.19)}} \leq \alpha - \nu - \gamma.
\end{aligned}$$

For $m \geq 1$ and $k = 1$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m,1) - v_{n+1}^{-}(m,0) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,1) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,0) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,2) + \frac{\gamma}{\alpha} \cdot v_n^{-}(m,0) \\
&\quad + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^{-}(m,1)\Big] \\
&\quad - \Big[0 + \frac{\nu}{\alpha} \cdot v_n^{-}(m,1) + \frac{\lambda+\mu+\gamma \cdot b}{\alpha} \cdot v_n^{-}(m,0)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^{-}(m+1,1) - v_n^{-}(m,1))}_{\leq \alpha, \quad \text{by (D.2.18)}} + \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^{-}(m,1) - v_n^{-}(m,0))}_{\leq \alpha, \quad \text{by (D.2.19)}} + \frac{\nu}{\alpha} \cdot \underbrace{(v_n^{-}(m,2) - v_n^{-}(m,1))}_{\leq \alpha, \quad \text{by (D.2.19)}} \\
&\quad + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v_n^{-}(m,1) - v_n^{-}(m,0))}_{\leq \alpha, \quad \text{by (D.2.19)}} + \underbrace{\mu - \frac{\mu}{\alpha} \cdot \underbrace{(v_n^{-}(m,0) - v_n^{-}(m-1,0))}_{\in [0,\alpha], \quad \text{by (D.2.17), (D.2.18)}}}_{\leq \mu} \\
&\leq 2\lambda + \nu + \mu + \gamma(b-1) \overset{(\lambda \leq \gamma)}{\leq} \alpha.
\end{aligned}$$

For $m \geq 1$ and $k \in \{2, \dots, b-1\}$ holds

$$\begin{aligned}
&v_{n+1}^{-}(m,k) - v_{n+1}^{-}(m,k-1) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,k) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,k-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,k+1) \\
&\quad + \frac{\gamma \cdot k}{\alpha} \cdot v_n^{-}(m,k-1) + \frac{\gamma \cdot (b-k)}{\alpha} \cdot v_n^{-}(m,k)\Big] \\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,k-1) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,k-2) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,k) \\
&\quad + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v_n^{-}(m,k-2) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v_n^{-}(m,k-1)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^{-}(m+1,k) - v_n^{-}(m+1,k-1))}_{\leq \alpha, \quad \text{by (D.2.19)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v_n^{-}(m-1,k-1) - v_n^{-}(m-1,k-2))}_{\leq \alpha, \quad \text{by (D.2.19)}} \\
&\quad + \frac{\nu}{\alpha} \cdot \underbrace{(v_n^{-}(m,k+1) - v_n^{-}(m,k))}_{\leq \alpha, \quad \text{by (D.2.19)}} + \frac{\gamma \cdot (k-1)}{\alpha} \cdot \underbrace{(v_n^{-}(m,k-1) - v_n^{-}(m,k-2))}_{\leq \alpha, \quad \text{by (D.2.19)}} \\
&\quad + \frac{\gamma \cdot (b-k)}{\alpha} \cdot \underbrace{(v_n^{-}(m,k) - v_n^{-}(m,k-1))}_{\leq \alpha, \quad \text{by (D.2.19)}} \leq \alpha - \gamma.
\end{aligned}$$

For $m \geq 1$ and $k = b$ holds

$$
\begin{aligned}
& v_{n+1}^{-}(m,b) - v_{n+1}^{-}(m,b-1) \\
&= \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,b) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,b-1) + \frac{\gamma \cdot b}{\alpha} \cdot v_n^{-}(m,b-1) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,b)\right] \\
&\quad - \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^{-}(m+1,b-1) + \frac{\mu}{\alpha} \cdot v_n^{-}(m-1,b-2) + \frac{\nu}{\alpha} \cdot v_n^{-}(m,b) \right. \\
&\quad \left. + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^{-}(m,b-2) + \frac{\gamma}{\alpha} \cdot v_n^{-}(m,b-1)\right] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v_n^{-}(m+1,b) - v_n^{-}(m+1,b-1))}_{\leq \alpha, \ \text{by (D.2.19)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v_n^{-}(m-1,b-1) - v_n^{-}(m-1,b-2))}_{\leq \alpha, \ \text{by (D.2.19)}} \\
&\quad + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v_n^{-}(m,b-1) - v_n^{-}(m,b-2))}_{\leq \alpha, \ \text{by (D.2.19)}} \leq (\lambda + \mu + \gamma \cdot (b-1)) = \alpha - \nu - \gamma.
\end{aligned}
$$

□

The following proposition is Proposition 4.3.13.

Proposition. *If in the production-inventory system with ageing regime* $k \to (\gamma \cdot k) \cdot 1_{\{m=0\}} + (\gamma \cdot (k-1)_+) \cdot 1_{\{nm>0\}}$ *we have* $\mu = \gamma$*, then the finite time cumulative rewards* $v_n^o(m,k)$ *are isotone with respect to the natural order.*

Proof. We show by induction isotonicity in both arguments, that the increase is bounded, and that v_n^o is concave in the time parameter m. For all $n \in \mathbb{N}$ holds

$$
\begin{aligned}
v_n^o(m,k) - v_n^o(m,k-1) &\geq 0, & \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0, \quad & \text{(D.2.20)} \\
v_n^o(m+1,k) - v_n^o(m,k) &\geq 0, & \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0, \quad & \text{(D.2.21)} \\
v_n^o(m+1,k) - v_n^o(m,k) &\leq \alpha, & \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0, \quad & \text{(D.2.22)} \\
v_n^o(m,k) - v_n^o(m,k-1) &\leq \alpha, & \forall k \in \{1,\dots,b\},\ m \in \mathbb{N}_0, \quad & \text{(D.2.23)} \\
v_n^o(m+1,k) - 2 \cdot v_n^o(m,k) + v_n^o(m-1,k) &\leq 0, & \forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}. \quad & \text{(D.2.24)}
\end{aligned}
$$

For $n = 1$ we have $v_1^o(m,k) = r(m,k) = \mu \cdot 1_{\{m>0\}} \cdot 1_{\{k>0\}}$, $\forall k \in \{0,1,\dots,b\},\ m \in \mathbb{N}_0$, so (D.2.20)-(D.2.23) are trivially true.

Assume that (D.2.20)-(D.2.24) hold for some $n \in \mathbb{N}$. We shall verify these properties for v_{n+1}^o. In any case we exploit again: $v_{n+1}^o = r + R^o \cdot v_n^o$, $n \geq 1$.

► First we check (D.2.20).
For $m = 0$ the proof of (D.2.20) is similar to that of (D.2.16).

For $m \geq 1$ and $k = 1$ holds

$$\begin{aligned}
& v^o_{n+1}(m,1) - v^o_{n+1}(m,0) \\
= & \left[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,1) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,0) + \frac{\nu}{\alpha} \cdot v^o_n(m,2) + \frac{\gamma \cdot b}{\alpha} \cdot v^o_n(m,1)\right] \\
& - \left[0 + \frac{\nu}{\alpha} \cdot v^o_n(m,1) + \frac{\lambda + \mu + \gamma \cdot b}{\alpha} \cdot v^o_n(m,0)\right] \\
= & \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+1,1) - v^o_n(m,1))}_{\geq 0, \quad \text{by (D.2.21)}} + \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m,1) - v^o_n(m,0))}_{\geq 0, \quad \text{by (D.2.20)}} + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m,2) - v^o_n(m,1))}_{\geq 0, \quad v\text{(D.2.20)}} \\
& + \frac{\gamma \cdot b}{\alpha} \cdot \underbrace{(v^o_n(m,1) - v^o_n(m,0))}_{\geq 0, \quad \text{by (D.2.20)}} + \underbrace{\mu - \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m,0) - v^o_n(m-1,0))}_{\in[0,\alpha], \quad \text{by (D.2.21), (D.2.22)}}}_{\geq 0} \geq 0.
\end{aligned}$$

For $m \geq 1$ and $k \in \{2, \ldots, b-1\}$ holds

$$\begin{aligned}
& v^o_{n+1}(m,k) - v^o_{n+1}(m,k-1) \\
= & \left[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,k) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m,k+1)\right. \\
& \left. + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(m,k)\right] \\
& - \left[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,k-1) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,k-2) + \frac{\nu}{\alpha} \cdot v^o_n(m,k)\right. \\
& \left. + \frac{\gamma \cdot (k-2)}{\alpha} \cdot v^o_n(m,k-2) + \frac{\gamma \cdot (b-k+2)}{\alpha} \cdot v^o_n(m,k-1)\right] \\
= & \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+1,k) - v^o_n(m+1,k-1))}_{\geq 0, \quad \text{by (D.2.20)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m-1,k-1) - v^o_n(m-1,k-2))}_{\geq 0, \quad \text{by (D.2.20)}} \\
& + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m,k+1) - v^o_n(m,k))}_{\geq 0, \quad \text{by (D.2.20)}} + \frac{\gamma \cdot (k-2)}{\alpha} \cdot \underbrace{(v^o_n(m,k-1) - v^o_n(m,k-2))}_{\geq 0, \quad \text{by (D.2.20)}} \\
& + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot \underbrace{(v^o_n(m,k) - v^o_n(m,k-1))}_{\leq \alpha, \quad \text{by (D.2.20)}} \geq 0.
\end{aligned}$$

For $m \geq 1$ and $k = b$ holds

$$
\begin{aligned}
& v^o_{n+1}(m,b) - v^o_{n+1}(m,b-1) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,b) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m,b-1) \\
&\quad + \frac{\nu+\gamma}{\alpha} \cdot v^o_n(m,b)\Big] \\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,b-1) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,b-2) + \frac{\nu}{\alpha} \cdot v^o_n(m,b) \\
&\quad + \frac{\gamma \cdot (b-2)}{\alpha} \cdot v^o_n(m,b-2) + \frac{\gamma \cdot 2}{\alpha} \cdot v^o_n(m,b-1)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+1,b) - v^o_n(m+1,b-1))}_{\geq 0, \text{ by (D.2.20)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m-1,b-1) - v^o_n(m-1,b-2))}_{\geq 0, \text{ by (D.2.20)}} \\
&\quad + \frac{\gamma \cdot (b-2)}{\alpha} \cdot \underbrace{(v^o_n(m,b-1) - v^o_n(m,b-2))}_{\geq 0, \text{ by (D.2.20)}} + \frac{\gamma}{\alpha} \cdot \underbrace{(v^o_n(m,b) - v^o_n(m,b-1))}_{\geq 0, \text{ by (D.2.20)}} \geq 0.
\end{aligned}
$$

► Second we check (D.2.21).
For $m = 0$ the proof of (D.2.21) is similar to that of (D.2.17).

For $m \geq 1$ and $k \in \{1, \ldots, b-1\}$ holds

$$
\begin{aligned}
& v^o_{n+1}(m+1,k) - v^o_{n+1}(m,k) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+2,k) + \frac{\mu}{\alpha} \cdot v^o_n(m,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m+1,k+1) \\
&\quad + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m+1,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(m+1,k)\Big] \\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,k) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m,k+1) \\
&\quad + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(m,k)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+2,k) - v^o_n(m+1,k))}_{\geq 0, \text{ by (D.2.21)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m,k-1) - v^o_n(m-1,k-1))}_{\geq 0, \text{ by (D.2.21)}} \\
&\quad + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m+1,k+1) - v^o_n(m,k+1))}_{\geq 0, \text{ by (D.2.21)}} + \frac{\gamma \cdot (k-1)}{\alpha} \cdot \underbrace{(v^o_n(m+1,k-1) - v^o_n(m,k-1))}_{\geq 0, \text{ by (D.2.21)}} \\
&\quad + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot \underbrace{(v^o_n(m+1,k) - v^o_n(m,k))}_{\geq 0, \text{ by (D.2.21)}} \geq 0.
\end{aligned}
$$

For $m \geq 1$ and $k = b$ holds

$$\begin{aligned}
&v^o_{n+1}(m+1,b) - v^o_{n+1}(m,b)\\
&= \Big[\mu + \frac{\lambda}{\alpha}\cdot v^o_n(m+2,b) + \frac{\mu}{\alpha}\cdot v^o_n(m,b-1) + \frac{\gamma\cdot(b-1)}{\alpha}\cdot v^o_n(m+1,b-1)\\
&\quad + \frac{\nu+\gamma}{\alpha}\cdot v^o_n(m+1,b)\Big]\\
&\quad - \Big[\mu + \frac{\lambda}{\alpha}\cdot v^o_n(m+1,b) + \frac{\mu}{\alpha}\cdot v^o_n(m-1,b-1) + \frac{\gamma\cdot(b-1)}{\alpha}\cdot v^o_n(m,b-1)\\
&\quad + \frac{\nu+\gamma}{\alpha}\cdot v^o_n(m,b)\Big]\\
&= \frac{\lambda}{\alpha}\cdot\underbrace{(v^o_n(m+2,b) - v^o_n(m+1,b))}_{\geq 0,\ \text{by (D.2.21)}} + \frac{\mu}{\alpha}\cdot\underbrace{(v^o_n(m,b-1) - v^o_n(m-1,b-1)}_{\geq 0,\ \text{by (D.2.21)}}\\
&\quad + \frac{\gamma\cdot(b-1)}{\alpha}\cdot\underbrace{(v^o_n(m+1,b-1) - v^o_n(m,b-1))}_{\geq 0,\ \text{by (D.2.21)}} + \frac{\nu+\gamma}{\alpha}\cdot\underbrace{(v^o_n(m+1,b) - v^o_n(m,b))}_{\geq 0,\ \text{by (D.2.21)}} \geq 0.
\end{aligned}$$

► Third we check (D.2.22).
For $m \geq 0$ and $k = 0$ holds

$$\begin{aligned}
&v^o_{n+1}(m+1,0) - v^o_{n+1}(m,0)\\
&= \Big[\frac{\nu}{\alpha}\cdot v^o_n(m+1,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v^o_n(m+1,0)\Big]\\
&\quad - \Big[\frac{\nu}{\alpha}\cdot v^o_n(m,1) + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot v^o_n(m,0)\Big]\\
&= \frac{\nu}{\alpha}\cdot\underbrace{(v^o_n(m+1,1) - v^o_n(m,1))}_{\leq\alpha,\ \text{by (D.2.22)}} + \frac{\lambda+\mu+\gamma\cdot b}{\alpha}\cdot\underbrace{(v^o_n(m+1,0) - v^o_n(m,0))}_{\geq\leq\alpha,\ \text{by (D.2.22)}} \leq \alpha.
\end{aligned}$$

For $m = 0$ and $k \in \{1,\ldots,b-1\}$ holds

$$\begin{aligned}
&v^o_{n+1}(1,k) - v^o_{n+1}(0,k)\\
&= \Big[\mu + \frac{\lambda}{\alpha}\cdot v^o_n(2,k) + \frac{\mu}{\alpha}\cdot v^o_n(0,k-1) + \frac{\nu}{\alpha}\cdot v^o_n(1,k+1)\\
&\quad + \frac{\gamma\cdot(k-1)}{\alpha}\cdot v^o_n(1,k-1) + \frac{\gamma\cdot(b-k+1)}{\alpha}\cdot v^o_n(1,k)\Big]\\
&\quad - \Big[0 + \frac{\lambda}{\alpha}\cdot v^o_n(1,k) + \frac{\nu}{\alpha}\cdot v^o_n(0,k+1) + \frac{\gamma\cdot k}{\alpha}\cdot v^o_n(0,k-1) + \frac{\mu+\gamma\cdot(b-k)}{\alpha}\cdot v^o_n(0,k)\Big]\\
&= \frac{\lambda}{\alpha}\cdot\underbrace{(v^o_n(2,k) - v^o_n(1,k))}_{\leq\alpha,\ \text{by (D.2.22)}} + \frac{\nu}{\alpha}\cdot\underbrace{(v^o_n(1,k+1) - v^o_n(0,k+1))}_{\leq\alpha,\ \text{by (D.2.22)}}\\
&\quad + \frac{\gamma\cdot(k-1)}{\alpha}\cdot\underbrace{(v^o_n(1,k-1) - v^o_n(0,k-1))}_{\leq\alpha,\ \text{by (D.2.22)}} + \frac{\gamma\cdot(b-k+1)}{\alpha}\cdot\underbrace{(v^o_n(1,k) - v^o_n(0,k))}_{\leq\alpha,\ \text{by (D.2.22)}}\\
&\quad + \mu + \underbrace{(\frac{\gamma}{\alpha} - \frac{\mu}{\alpha})}_{=0,\ \text{by } (\gamma=\mu)}\cdot(v^o_n(0,k) - v^o_n(0,k-1)) \leq \alpha.
\end{aligned}$$

For $m = 0$ and $k = b$ holds

$$
\begin{aligned}
& v^o_{n+1}(1,b) - v^o_{n+1}(0,b) \\
&= \left[\mu + \frac{\lambda}{\alpha}\cdot v^o_n(2,b) + \frac{\mu}{\alpha}\cdot v^o_n(0,b-1) + \frac{\gamma\cdot(b-1)}{\alpha}\cdot v^o_n(1,b-1) + \frac{\nu+\gamma}{\alpha}\cdot v^o_n(1,b)\right] \\
&\quad - \left[0 + \frac{\lambda}{\alpha}\cdot v^o_n(1,b) + \frac{\gamma\cdot b}{\alpha}\cdot v^o_n(0,b-1) + \frac{\mu+\nu}{\alpha}\cdot v^o_n(0,b)\right] \\
&= \frac{\lambda}{\alpha}\cdot \underbrace{(v^o_n(2,b) - v^o_n(1,b))}_{\leq\alpha,\ \text{ by (D.2.22)}} + \frac{\gamma\cdot(b-1)}{\alpha}\cdot \underbrace{(v^o_n(1,b-1) - v^o_n(0,b-1))}_{\leq\alpha,\ \text{ by (D.2.22)}} \\
&\quad + \frac{\nu+\gamma}{\alpha}\cdot \underbrace{(v^o_n(1,b) - v^o_n(0,b))}_{\leq\alpha,\ \text{ by (D.2.22)}} + \mu + \underbrace{(\frac{\gamma}{\alpha} - \frac{\mu}{\alpha})}_{=0,\ \text{by } (\gamma=\mu)} \cdot (v^o_n(0,b) - v^o_n(0,b-1)) \leq \alpha.
\end{aligned}
$$

For $m \geq 1$ and $k \in \{1, \dots, b-1\}$ holds

$$
\begin{aligned}
& v^o_{n+1}(m+1,k) - v^o_{n+1}(m,k) \\
&= \Bigg[\mu + \frac{\lambda}{\alpha}\cdot v^o_n(m+2,k) + \frac{\mu}{\alpha}\cdot v^o_n(m,k-1) + \frac{\nu}{\alpha}\cdot v^o_n(m+1,k+1) \\
&\quad + \frac{\gamma\cdot(k-1)}{\alpha}\cdot v^o_n(m+1,k-1) + \frac{\gamma\cdot(b-k+1)}{\alpha}\cdot v^o_n(m+1,k)\Bigg] \\
&\quad - \Bigg[\mu + \frac{\lambda}{\alpha}\cdot v^o_n(m+1,k) + \frac{\mu}{\alpha}\cdot v^o_n(m-1,k-1) + \frac{\nu}{\alpha}\cdot v^o_n(m,k+1) \\
&\quad + \frac{\gamma\cdot(k-1)}{\alpha}\cdot v^o_n(m,k-1) + \frac{\gamma\cdot(b-k+1)}{\alpha}\cdot v^o_n(m,k)\Bigg] \\
&= \frac{\lambda}{\alpha}\cdot \underbrace{(v^o_n(m+2,k) - v^o_n(m+1,k))}_{\leq\alpha,\ \text{ by (D.2.22)}} + \frac{\mu}{\alpha}\cdot \underbrace{(v^o_n(m,k-1) - v^o_n(m-1,k-1))}_{\leq\alpha,\ \text{ by (D.2.22)}} \\
&\quad + \frac{\nu}{\alpha}\cdot \underbrace{(v^o_n(m+1,k+1) - v^o_n(m,k+1))}_{\leq\alpha,\ \text{ by (D.2.22)}} + \frac{\gamma\cdot(k-1)}{\alpha}\cdot \underbrace{(v^o_n(m+1,k-1) - v^o_n(m,k-1))}_{\leq\alpha,\ \text{ by (D.2.22)}} \\
&\quad + \frac{\gamma\cdot(b-k+1)}{\alpha}\cdot \underbrace{(v^o_n(m+1,k) - v^o_n(m,k))}_{\leq\alpha,\ \text{ by (D.2.22)}} \leq \alpha.
\end{aligned}
$$

For $m \geq 1$ and $k = b$ holds

$$\begin{aligned}
& v^o_{n+1}(m+1,b) - v^o_{n+1}(m,b) \\
= & \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+2,b) + \frac{\mu}{\alpha} \cdot v^o_n(m,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m+1,b-1) \\
& + \frac{\nu+\gamma}{\alpha} \cdot v^o_n(m+1,b)\Big] \\
& - \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,b) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m,b-1) \\
& + \frac{\nu+\gamma}{\alpha} \cdot v^o_n(m,b)\Big] \\
= & \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+2,b) - v^o_n(m+1,b))}_{\leq \alpha, \text{ by (D.2.22)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m,b-1) - v^o_n(m-1,b-1))}_{\leq \alpha, \text{ by (D.2.23)}} \\
& + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{(v^o_n(m+1,b-1) - v^o_n(m,b-1))}_{\leq \alpha, \text{ by (D.2.22)}} + \frac{\nu+\gamma}{\alpha} \cdot \underbrace{(v^o_n(m+1,b) - v^o_n(m,b))}_{\leq \alpha, \text{ by (D.2.22)}} \leq \alpha.
\end{aligned}$$

► Fourth we check (D.2.23).
For $m = 0$ the proof of (D.2.23) is the same as of (D.2.19).

For $m \geq 1$ and $k = 1$ holds

$$\begin{aligned}
& v^o_{n+1}(m,1) - v^o_{n+1}(m,0) \\
= & \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,1) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,0) + \frac{\nu}{\alpha} \cdot v^o_n(m,2) + \frac{\gamma \cdot b}{\alpha} \cdot v^o_n(m,1)\Big] \\
& - \Big[0 + \frac{\nu}{\alpha} \cdot v^o_n(m,1) + \frac{\lambda+\mu+\gamma \cdot b}{\alpha} \cdot v^o_n(m,0)\Big] \\
= & \mu + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m,2) - v^o_n(m,1))}_{\leq \alpha, \text{ by (D.2.23)}} + \frac{\gamma \cdot b}{\alpha} \cdot \underbrace{(v^o_n(m,1) - v^o_n(m,0))}_{\leq \alpha, \text{ by (D.2.23)}} \\
& + \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+1,1) - v^o_n(m+1,0))}_{\leq \alpha, \text{ by (D.2.23)}} \\
& + \underbrace{\frac{\lambda}{\alpha} \cdot (v^o_n(m+1,0) - v^o_n(m,0)) - \frac{\mu}{\alpha} \cdot (v^o_n(m,0) - v^o_n(m-1,0))}_{\leq 0, \text{ by } \lambda<\mu \text{ and (D.2.21), (D.2.24)}} \leq \alpha.
\end{aligned}$$

For $m \geq 1$ and $k \in \{2, \ldots, b-1\}$ holds

$$
\begin{aligned}
& v^o_{n+1}(m,k) - v^o_{n+1}(m,k-1) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,k) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m,k+1) \\
&\quad + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(m,k)\Big] \\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,k-1) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,k-2) + \frac{\nu}{\alpha} \cdot v^o_n(m,k) \\
&\quad + \frac{\gamma \cdot (k-2)}{\alpha} \cdot v^o_n(m,k-2) + \frac{\gamma \cdot (b-k+2)}{\alpha} \cdot v^o_n(m,k-1)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+1,k) - v^o_n(m+1,k-1))}_{\leq \alpha, \text{ by (D.2.23)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m-1,k-1) - v^o_n(m-1,k-2))}_{\leq \alpha, \text{ by (D.2.23)}} \\
&\quad + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m,k+1) - v^o_n(m,k))}_{\leq \alpha, \text{ by (D.2.23)}} + \frac{\gamma \cdot (k-2)}{\alpha} \cdot \underbrace{(v^o_n(m,k-1) - v^o_n(m,k-2))}_{\leq \alpha, \text{ by (D.2.23)}} \\
&\quad + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot \underbrace{(v^o_n(m,k) - v^o_n(m,k-1))}_{\leq \alpha, \text{ by (D.2.23)}} \leq \alpha - \gamma.
\end{aligned}
$$

For $m \geq 1$ and $k = b$ holds

$$
\begin{aligned}
& v^o_{n+1}(m,b) - v^o_{n+1}(m,b-1) \\
&= \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,b) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m,b-1) \\
&\quad + \frac{\nu + \gamma}{\alpha} \cdot v^o_n(m,b)\Big] \\
&\quad - \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,b-1) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,b-2) + \frac{\nu}{\alpha} \cdot v^o_n(m,b) \\
&\quad + \frac{\gamma \cdot (b-2)}{\alpha} \cdot v^o_n(m,b-2) + \frac{\gamma \cdot 2}{\alpha} \cdot v^o_n(m,b-1)\Big] \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+1,b) - v^o_n(m+1,b-1))}_{\leq \alpha, \text{ by (D.2.23)}} + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m-1,b-1) - v^o_n(m-1,b-2))}_{\leq \alpha, \text{ by (D.2.23)}} \\
&\quad + \frac{\gamma \cdot (b-2)}{\alpha} \cdot \underbrace{(v^o_n(m,b-1) - v^o_n(m,b-2))}_{\leq \alpha, \text{ by (D.2.23)}} + \frac{\gamma}{\alpha} \cdot \underbrace{(v^o_n(m,b) - v^o_n(m,b-1))}_{\leq \alpha, \text{ by (D.2.23)}} \\
&\leq (\lambda + \mu + \gamma \cdot (b-1)) = \alpha - \nu - \gamma.
\end{aligned}
$$

► Fifth we check (D.2.24).
For $m \geq 1$ and $k = 0$ holds

$$\begin{aligned}
& v^o_{n+1}(m+1,0) - 2 \cdot v^o_{n+1}(m,0) + v^o_{n+1}(m-1,0) \\
= & \left[\frac{\nu}{\alpha} \cdot v^o_n(m+1,1) + \frac{\lambda+\mu+\gamma \cdot b}{\alpha} \cdot v^o_n(m+1,0)\right] \\
& -2 \cdot \left[\frac{\nu}{\alpha} \cdot v^o_n(m,1) + \frac{\lambda+\mu+\gamma \cdot b}{\alpha} \cdot v^o_n(m,0)\right] \\
& + \left[\frac{\nu}{\alpha} \cdot v^o_n(m-1,1) + \frac{\lambda+\mu+\gamma \cdot b}{\alpha} \cdot v^o_n(m-1,0)\right] \\
= & \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m+1,1) - 2 \cdot v^o_n(m,1) + v^o_n(m-1,1))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\lambda+\mu+\gamma \cdot b}{\alpha} \cdot \underbrace{(v^o_n(m+1,0) - 2 \cdot v^o_n(m,0) + v^o_n(m-1,0))}_{\leq 0, \quad \text{by (D.2.24)}} \leq 0.
\end{aligned}$$

For $m = 1$ and $k \in \{1, \ldots, b-1\}$ holds

$$\begin{aligned}
& v^o_{n+1}(2,k) - 2 \cdot v^o_{n+1}(1,k) + v^o_{n+1}(0,k) \\
= & \left[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(3,k) + \frac{\mu}{\alpha} \cdot v^o_n(1,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(2,k+1) + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(2,k-1)\right. \\
& \left. + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(2,k)\right] \\
& -2 \cdot \left[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(2,k) + \frac{\mu}{\alpha} \cdot v^o_n(0,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(1,k+1)\right. \\
& \left. + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(1,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(1,k)\right] \\
& + \left[0 + \frac{\lambda}{\alpha} \cdot v^o_n(1,k) + \frac{\nu}{\alpha} \cdot v^o_n(0,k+1) + \frac{\gamma \cdot k}{\alpha} \cdot v^o_n(0,k-1) + \frac{\mu + \gamma(b-k)}{\alpha} \cdot v^o_n(0,k)\right] \\
= & -\mu + \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(3,k) - 2 \cdot v^o_n(2,k) + v^o_n(1,k))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(2,k+1) - 2 \cdot v^o_n(1,k+1) + v^o_n(0,k+1))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\gamma(k-1)}{\alpha} \cdot \underbrace{(v^o_n(2,k-1) - 2 \cdot v^o_n(1,k-1) + v^o_n(0,k-1))}_{\leq 0, \quad \text{by (D.2.24)}} + \frac{\gamma}{\alpha} \cdot v^o_n(0,k-1) \\
& + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot \underbrace{(v^o_n(2,k) - 2 \cdot v^o_n(1,k) + v^o_n(0,k))}_{\leq 0, \quad \text{by (D.2.24)}} - \frac{\gamma}{\alpha} \cdot v^o_n(0,k) \\
& + \frac{\mu}{\alpha} \cdot (v^o_n(1,k-1) - 2 \cdot v^o_n(0,k-1) + v^o_n(0,k))
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^o(3,k) - 2 \cdot v_n^o(2,k) + v_n^o(1,k)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^o(2,k+1) - 2 \cdot v_n^o(1,k+1) + v_n^o(0,k+1)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\gamma(k-1)}{\alpha} \cdot \underbrace{\left(v_n^o(2,k-1) - 2 \cdot v_n^o(1,k-1) + v_n^o(0,k-1)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot \underbrace{\left(v_n^o(2,k) - 2 \cdot (v_n^o(1,k) + v_n^o(0,k)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\mu}{\alpha} \cdot \underbrace{\left(v_n^o(1,k-1) - v_n^o(0,k-1)\right)}_{\in [0,\alpha] \quad \text{by (D.2.21), (D.2.22)}} - \mu \\
&\quad + \underbrace{\frac{\mu}{\alpha} \cdot \left(v_n^o(0,k) - v_n^o(0,k-1)\right) - \frac{\gamma}{\alpha} \cdot \left(v_n^o(0,k) - v_n^o(0,k-1)\right)}_{=0, \quad \text{by } \gamma = \mu} \leq 0.
\end{aligned}
$$

For $m = 1$ and $k = b$ holds

$$
\begin{aligned}
&v_{n+1}^o(2,b) - 2 \cdot v_{n+1}^o(1,b) + v_{n+1}^o(0,b) \\
&= \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^o(3,b) + \frac{\mu}{\alpha} \cdot v_n^o(1,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^o(2,b-1) + \frac{\nu + \gamma}{\alpha} \cdot v_n^o(2,b)\right] \\
&\quad - 2 \cdot \left[\mu + \frac{\lambda}{\alpha} \cdot v_n^o(2,b) + \frac{\mu}{\alpha} \cdot v_n^o(0,b-1) + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v_n^o(1,b-1) + \frac{\nu + \gamma}{\alpha} \cdot v_n^o(1,b)\right] \\
&\quad + \left[0 + \frac{\lambda}{\alpha} \cdot v_n^o(1,b) + \frac{\gamma \cdot b}{\alpha} \cdot v_n^o(0,b-1) + \frac{\mu + \nu}{\alpha} \cdot v_n^o(0,b)\right] \\
&= -\mu + \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^o(3,b) - 2 \cdot v_n^o(2,b) + v_n^o(1,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^o(2,b) - 2 \cdot v_n^o(1,b) + v_n^o(0,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{\left(v_n^o(2,b-1) - 2 \cdot v_n^o(1,b-1) + v_n^o(0,b-1)\right)}_{\leq 0, \quad \text{by (D.2.24)}} + \frac{\gamma}{\alpha} \cdot v_n^o(0,b-1) \\
&\quad + \frac{\gamma}{\alpha} \cdot \underbrace{\left(v_n^o(2,b) - 2 \cdot v_n^o(1,b) + v_n^o(0,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} - \frac{\gamma}{\alpha} \cdot v_n^o(0,b) \\
&\quad + \frac{\mu}{\alpha} \cdot \left(v_n^o(1,b-1) - 2 \cdot v_n^o(0,b-1) + v_n^o(0,b)\right) \\
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^o(3,b) - 2 \cdot v_n^o(2,b) + v_n^o(1,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} + \frac{\nu}{\alpha} \cdot \underbrace{\left(v_n^o(2,b) - 2 \cdot v_n^o(1,b) + v_n^o(0,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{\left(v_n^o(2,b-1) - 2 \cdot v_n^o(1,b-1) + v_n^o(0,b-1)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\gamma}{\alpha} \cdot \underbrace{\left(v_n^o(2,b) - 2 \cdot v_n^o(1,b) + v_n^o(0,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \underbrace{- \mu + \frac{\mu}{\alpha} \cdot \underbrace{\left(v_n^o(1,b-1) - v_n^o(0,b-1)\right)}_{\in [0,\alpha] \quad \text{by (D.2.21),(D.2.22)}}}_{\leq 0} \\
&\quad + \underbrace{\frac{\mu}{\alpha} \cdot \left(v_n^o(0,b) - v_n^o(0,b-1)\right) - \frac{\gamma}{\alpha} \cdot \left(v_n^o(0,b) - v_n^o(0,b-1)\right)}_{=0, \quad \text{by } \gamma = \mu} \leq 0.
\end{aligned}
$$

For $m \geq 2$ and $k \in \{1, \ldots, b-1\}$ holds

$$
\begin{aligned}
& v^o_{n+1}(m+1,k) - 2 \cdot v^o_{n+1}(m,k) + v^o_{n+1}(m-1,k) \\
= & \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+2,k) + \frac{\mu}{\alpha} \cdot v^o_n(m,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m+1,k+1) \\
& + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m+1,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(m+1,k)\Big] \\
& -2 \cdot \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,k) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m,k+1) \\
& + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m,k-1) + \frac{\gamma(b-k+1)}{\alpha} \cdot v^o_n(m,k)\Big] \\
& + \Big[[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m,k) + \frac{\mu}{\alpha} \cdot v^o_n(m-2,k-1) + \frac{\nu}{\alpha} \cdot v^o_n(m-1,k+1) \\
& + \frac{\gamma \cdot (k-1)}{\alpha} \cdot v^o_n(m-1,k-1) + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot v^o_n(m-1,k)\Big] \\
= & \frac{\lambda}{\alpha} \cdot \underbrace{(v^o_n(m+2,k) - 2 \cdot v^o_n(m+1,k) + v^o_n(m,k))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\mu}{\alpha} \cdot \underbrace{(v^o_n(m,k-1) - 2 \cdot v^o_n(m-1,k-1) + v^o_n(m-2,k-1))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\nu}{\alpha} \cdot \underbrace{(v^o_n(m+1,k+1) - 2 \cdot v^o_n(m,k+1) + v^o_n(m-1,k+1))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\gamma \cdot (k-1)}{\alpha} \cdot \underbrace{(v^o_n(m+1,k-1) - 2 \cdot v^o_n(m,k-1) + v^o_n(m-1,k-1))}_{\leq 0, \quad \text{by (D.2.24)}} \\
& + \frac{\gamma \cdot (b-k+1)}{\alpha} \cdot \underbrace{(v^o_n(m+1,k) - 2 \cdot v^o_n(m,k) + v^o_n(m-1,k))}_{\leq 0, \quad \text{by (D.2.24)}} \leq 0.
\end{aligned}
$$

For $m \geq 2$ and $k = b$ holds

$$
\begin{aligned}
& v^o_{n+1}(m+1,b) - 2 \cdot v^o_{n+1}(m,b) + v^o_{n+1}(m-1,b) \\
= & \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+2,b) + \frac{\mu}{\alpha} \cdot v^o_n(m,b-1) \\
& + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m+1,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v^o_n(m+1,b)\Big] \\
& -2 \cdot \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m+1,b) + \frac{\mu}{\alpha} \cdot v^o_n(m-1,b-1) \\
& + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v^o_n(m,b)\Big] \\
& + \Big[\mu + \frac{\lambda}{\alpha} \cdot v^o_n(m,b) + \frac{\mu}{\alpha} \cdot v^o_n(m-2,b-1) \\
& + \frac{\gamma \cdot (b-1)}{\alpha} \cdot v^o_n(m-1,b-1) + \frac{\nu+\gamma}{\alpha} \cdot v^o_n(m-1,b)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\lambda}{\alpha} \cdot \underbrace{\left(v_n^o(m+2,b) - 2 \cdot v_n^o(m+1,b) + v_n^o(m,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\mu}{\alpha} \cdot \underbrace{\left(v_n^o(m,b-1) - 2 \cdot v_n^o(m-1,b-1) + v_n^o(m-2,b-1)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\gamma \cdot (b-1)}{\alpha} \cdot \underbrace{\left(v_n^o(m+1,b-1) - 2 \cdot v_n^o(m,b-1) + v_n^o(m-1,b-1)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \\
&\quad + \frac{\nu + \gamma}{\alpha} \cdot \underbrace{\left(v_n^o(m+1,b) - 2 \cdot v_n^o(m,b) + v_n^o(m-1,b)\right)}_{\leq 0, \quad \text{by (D.2.24)}} \leq 0.
\end{aligned}
$$

□

E. Appendix to Chapter 11

E.1. Proof of irreducibility

The stochastic queueing-inventory process Z as described in Chapter 11 is irreducible because it can be shown that the state $(0, 0, 0)$ can be reached from any other state and vice versa. We sketch the arguments for the case $p = 1$.

Firstly, we will show that the state $(0, 0, 0)$ can be reached from any other state (n_1, n_2, k) with $n_1 \geq 0$, $n_2 \geq 0$ and $0 \leq k \leq b$. The inter-arrival times of raw material at the inventory can be arbitrarily small, so that the replenishment of raw material at the inventory is very fast. Therefore, the inventory becomes greater than zero so that after a finite number of transitions the priority queue will become zero. The ordinary customers can be served only if no priority customer is present. Due to the fact that the priority queue will become zero after a finite number of transitions and that the inter-arrival times of the priority customers can be arbitrarily large and the service times can be arbitrarily small, the ordinary queue can also become zero after a finite number of transitions. Since the number of services in a busy period of the server[1] can be any positive number with positive probability, it is possible that at the end of the described busy period the inventory level becomes zero.

Secondly, we will show that the states (n_1, n_2, k) with $n_1 \geq 0$, $n_2 \geq 0$ and $0 < k \leq b$ can be reached from $(0, 0, 0)$. The inter-arrival times of the raw material at the inventory can be arbitrarily small, so that the replenishment of the raw material at the inventory is very fast. Thus, it is possible that the inventory level becomes k. Then, the inter-arrival times of ordinary and priority customers can be arbitrarily small. Hence, it is possible that n_1 resp. n_2 customers are in the priority queue resp. in the ordinary queue.

Thirdly, we will show that the states $(n_1, n_2, 0)$ with $n_1 \geq 0$ and $n_2 \geq 0$ can be reached from $(0, 0, 0)$. The inter-arrival times of raw material at the inventory can be arbitrarily small, so that the replenishment of raw material at the inventory is very fast. Thus, it is possible that the inventory level becomes m. Then, the inter-arrival times of ordinary and priority customers can be arbitrarily small. Hence, it is possible that $n_1 + m$ resp. n_2 customers are in the priority queue resp. in the ordinary queue, $m \in \mathbb{N}$. Since the number of services in a busy period of the server can be any positive number with positive probability, it is possible that at the end of a busy period the inventory level becomes zero and that n_1 resp. n_2 customers are in the priority queue resp. in the ordinary queue.

[1] A busy period is defined as a time period during which the server serves without interruption customers, i.e. a time period during which at least one customer is present <u>and</u> the inventory is not empty. Hence, a busy period begins when an idling server restarts serving some customer and ends when the inventory or the queue at the production system becomes empty next.

List of abbreviations

cf.	(confer) compare with
e.g.	(exempli gratia) for example
et al.	(et alii) and others
f., ff.	and the following page (f.) or pages (ff.)
FCFS	first-come, first-served
GBE	global balance equation
i.e.	(id est) that is
LCFS-PR	last-come, first-served preemptive resume
LDQBD	level-dependent quasi-birth-and-death
p., pp.	page(s)
resp.	respectively

Index

Bibliography

[AGR07] H. Arslan, S.C. Graves, and T.A. Roemer. A single-product inventory model for multiple demand classes. *Management Science*, 53(9):1486–1500, 2007.

[AMBB14] H. Abouee-Mehrizi, O. Baron, and O. Berman. Exact Analysis of Capacitated Two-Echelon Inventory Systems with Priorities. *Manufacturing & Service Operations Management*, 16(4):561–577, 2014.

[ARS15] M. Amirthakodi, V. Radhamani, and B. Sivakumar. A perishable inventory system with service facility and feedback customers. *Annals of Operations Research*, 233:25–55, 2015.

[Asm03] S. Asmussen. *Applied Probability and Queues*, volume 51 of *Applications of Mathematics*. Springer, New York, second edition, 2003.

[Axs00] S. Axsäter. *Inventory Control*. Kluwer Academic Publishers, Norwell, Massachusetts, 2000.

[Bar10] O. Baron. Managing Perishable Inventory. In J.J. Cochran, editor, *Wiley Encyclopedia of Operations Research and Management Science*, pages 1–10. John Wiley & Sons, 2010.

[BCMP75] F. Baskett, M. Chandy, R. Muntz, and F.G. Palacios. Open, closed and mixed networks of queues with different classes of customers. *Journal of the Association for Computing Machinery*, 22:248–260, 1975.

[BCST09] D. Beyer, F. Cheng, S.P. Sethi, and M. Taksar. *Markovian Demand Inventory Models*. Springer US, 2009.

[BDO01] S. Balsamo, V. De Nitto Persone, and R. Onvural. *Analysis of Queueing Networks with Blocking*. Kluwer Academic Publisher, Norwell, 2001.

[BI83] S. Balsamo and G. Iazeolla. Aggregation and Disaggregation in Queueing Networks: The Principle of Product-Form Synthesis. In *Conference: Computer Performance and Reliability, Proceedings of the International Workshop, Pisa, Italy*, pages 111–125, 1983.

[BK99] O. Berman and E. Kim. Stochastic models for inventory management at service facilities. *Communications in Statistics. Stochastic Models*, 15(4):695–718, 1999.

[BK01] O. Berman and E. Kim. Dynamic order replenishment policy in internet-based supply chains. *Mathematical Methods of Operations Research*, 53:371–390, 2001.

[BK04] O. Berman and E. Kim. Dynamic inventory strategies for profit maximization in a service facility with stochastic service, demand and lead time. *Mathematical Methods of Operations Research*, 60:497–521, 2004.

[BL11] E. abd Stadje W. Bar-Lev, A.K. Schulte-Geers. On the demand pooling anomaly in inventory theory. *Operations Research Letters*, 39:411–413, 2011.

[BRT12] M. Bakker, J. Riezebos, and R.H. Teunter. Review of inventory systems with deterioration since 2001. *European Journal of Operational Research*, 221:275–284, 2012.

[BS00] O. Berman and K.P. Sapna. Inventory management at service facilities for systems with arbitrarily distributed service times. *Communications in Statistics. Stochastic Models*, 16(3&4):343–360, 2000.

[BS01] O. Berman and K.P. Sapna. Optimal control of service for facilities holding inventory. *Computers & Operations Research*, 28:429–441, 2001.

[BT95] L. Bright and P.G. Taylor. Calculating the equilibrium distribution in level dependent quasi-birth-and-death processes. *Communications in Statistics. Stochastic Models*, 11(3):497–525, 1995.

[BV11] M. Bijvank and F.A. Vis. Lost-sales inventory theory: A review. *European Journal of Operations Research*, 215:1–13, 2011.

[Chr98] M. Christopher. *Logistics and Supply Chain Management: Strategies for Reducing Cost and Improving Service*. Financial Times, London, 1998.

[CHW75] K.M. Chandy, U. Herzog, and L. Woo. Parametric Analysis of Queueing Networks. *IBM Journal of Research and Development*, 19:43–49, 1975.

[CK77] Y.C. Chow and W.H. Kohler. Dynamic load balancing in homogeneous two-processor distributed systems. In K.M. Chandy and M. Reiser, editors, *Computer Performance*, pages 39–52, North-Holland, Amsterdam, 1977.

[CK79] Y.C. Chow and W.H. Kohler. Models for dynamic load balancing in a heterogeneous multiple processor system. *IEEE Transactions on computers*, c-28(5):354–361, May 1979.

[CM96] R. Chakka and I. Mitrani. Approximate solutions for open networks with breakdowns and repairs. In F.P. Kelly, S. Zachary, and I. Ziedins, editors, *Stochastic Networks, Theory and Applications*, volume 4 of *Royal Statistical Society Lecture Notes Series*, chapter 16, pages 267–280. Clarendon Press, Oxford, 1996.

[CMP99] X. Chao, M. Miyazawa, and M. Pinedo. *Queueing Networks – Customers, Signals, and Product Form Solutions*. Wiley, Chichester, 1999.

[CZB01] C.Y. Chen, Z.Y. Zhao, and M.O. Ball. Quantity and Due Date Quoting Available to Promise. *Information Systems Frontiers*, 3(4):477–488, 2001.

[CZL12] H. Chen, Z. Zhou, and M. Liu. A queueing-inventory system with two classes demand and subject to selected service. *Journal of Information & Computational Science*, 9:3081–3089, 2012.

[Dad85] H. Daduna. The Cycle-Time Distribution in a Central Server Network with State-Dependent Branching. *Optimization*, 16:617–626, 1985.

[Dad87] H. Daduna. Cycle Times in a Starlike Network with State-Dependent Routing. *Journal of Applied Mathematics and Simulation*, 1:1–12, 1987.

[Dad90] H. Daduna. Exchangeable Items in Repair Systems: Delay Times. *Operations Research*, 38(2):349–354, 1990.

[Dad01a] H. Daduna. *Queueing Networks with Discrete Time Scale: Explicit Expressions for the Steady State Behavior of Discrete Time Stochastic Networks*, volume 2046 of *Lecture Notes in Computer Science*. Springer, Berlin, 2001.

[Dad01b] H. Daduna. Stochastic networks with product form equilibrium. In D.N. Shanbhag and C.R. Rao, editors, *Stochastic Processes: Theory and Methods*, volume 19 of *Handbook of Statistics*, chapter 11, pages 309–364. Elsevier Science, Amsterdam, 2001.

[DB13] K.H. van Donselaar and R.A.C.M. Broekmeulen. Determination of safety stocks in a lost sales inventory system with periodic review, positive lead-time, lot-sizing and a target fill rate. *International Journal of Production Economics*, 143:440–448, 2013.

[DCIM12] A. Di Crescenzo, A. Iuliano, and B. Martinucci. On a bilateral birth-death process with alternating rates. *Ricerche di Matematica*, 61(1):157–169, 2012.

[DCMM14] A. Di Crescenzo, C. Macci, and B. Martinucci. Asymptotic Results for Random Walks in Continuous Time with Alternating Rates. *Journal of Statistical Physics*, 154(5):1352–1364, 2014.

[Dij88] N.M. van Dijk. On Jackson's product form with 'jump-over' blocking. *Operations Research Letters*, 7(5):233–235, 1988.

[Dij93] N.M. van Dijk. *Queueing Networks and Product Forms – A Systems Approach*. Wiley, Chichester, 1993.

[Dij98] N.M. van Dijk. Bounds and error bounds for queueing networks. *Annals of Operations Research*, 79:295–319, 1998.

[Dij11a] N.M. van Dijk. Error Bounds and Comparison Results: The Markov Reward Approach for Queueing Networks. In R.J. Boucherie and N.M. van Dijk, editors, *Queueing Networks: A Fundamental Approach*, volume 154 of *International Series in Operations Research and Management Science*, chapter 9, pages 397–459. Springer, New York, 2011.

[Dij11b] N.M. van Dijk. On Practical Product Form Characterizations. In R.J. Boucherie and N.M. van Dijk, editors, *Queueing Networks: A Fundamental*

Approach, volume 154 of *International Series in Operations Research and Management Science*, chapter 1, pages 1–83. Springer, New York, 2011.

[DK92] N.M. van Dijk and H. Korezioglu. On product form approximations for communication networks with losses: error bounds. *Annals of Operations Research*, 35:69–94, 1992.

[DS96] H. Daduna and R. Szekli. A queueing theoretical proof of increasing property of Polya frequency functions. *Statistics and Probability Letters*, 26:233–242, 1996.

[DS09] N.M. van Dijk and E. van der Sluis. Pooling is not the answer. *European Journal of Operational Research*, 197(1):415–421, 2009.

[DW89] N.M. van Dijk and J. van Wal. Simple bounds and monotonicity results for multi-server exponential tandem queues. *Queueings Systems*, 4:1–16, 1989.

[Epp79] G.D. Eppen. Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science*, 25(5):498–501, 1979.

[FM77] L. Flatto and H.P. McKean. Two Queues in Parallel. *Communications on pure and applied mathematics*, pages 255–263, 1977.

[Fos53] F. G. Foster. On the stochastic matrices associated with certain queuing processes. *The Annals of Mathematical Statistics*, 24(3):355–360, 1953.

[FST12] S. Foss, S. Shneer, and A. Tyurlikov. Stability of a Markov-modulated Markov Chain, with Application to a Wireless Network Governed by two Protocols. *Stochastic Systems*, 2(1):208–231, 2012.

[GCB02] T.W. Gruen, D. Corsten, and S. Bharadwaj. Retail out-of-stocks: A world-wide examination of extent causes and consumer responses. *Grocery Manufacutres of America*, 2002.

[GG01] S.K. Goyal and B.C. Giri. Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*, 134:1–16, 2001.

[GK15] R. Gössinger and S. Kalkowski. Order Promising – A Robust Customer-Oriented Approach. In J. Dethloff, H.D. Haasis, H. Kopfer, H. Kotzab, and J. Schönberger, editors, *Logistics Management: Products, Actors, Technology – Proceedings of the German Academic Association for Business Research, Bremen, 2013*, Lecture Notes in Logistics, pages 135–149. Springer International Publishing Switzerland, 2015.

[GKAR12] J.P. van der Gaast, M.B.M. de Koster, I.J.B.F. Adan, and J.A.C. Resing. Modeling and performance analysis of sequential zone picking systems. Eurandom Preprint Series, November 2012.

[GN67] W.J. Gordon and G.F. Newell. Closed queueing networks with exponential servers. *Operations Research*, 15:254–265, 1967.

[HB11] T. Huisman and R.J. Boucherie. Discrete time networks with product form steady states. In R.J. Boucherie and N.M. van Dijk, editors, *Queueing Networks: A Fundamental Approach*, volume 154 of *International Series in Operations Research and Management Science*, chapter 7, pages 313–344. Springer, New York, 2011.

[HD84] A. Hordijk and N.M. van Dijk. Networks of Queues. Part I, II. In F. Baccelli and G. Fayolle, editors, *Modelling and Performance Evaluation Methodology. Lecture Notes in Control and Information Sciences*, volume 60, pages 151–205, Berlin – Heidelberg, 1984. Springer.

[Hen90] W. Henderson. Performance analysis: When do we give up on product form solutions. *Computer Networks and ISDN Systems*, 20(1):271–275, 1990.

[HHS11] R. Haji, A. Haji, and M. Saffari. Queueing Inventory System in a Two-level Supply Chain with One-for-One Ordering Policy. *Journal of Industrial and Systems Engineering*, 5(1):52–62, 2011.

[HJB02a] Q.-M. He, E. M. Jewkes, and J. Buzacott. Optimal and near-optimal inventory control policies for a make-to-order inventory-production system. *European Journal of Operational Research*, 141(1):113–132, 2002.

[HJB02b] Q.-M. He, E.M. Jewkes, and J. Buzacott. The value of information used in inventory control of a make-to-order inventory-production system. *IIE Transactions*, 34(11):999–1013, 2002.

[HRS17] K. Hinderer, U. Rieder, and M. Stieglitz. *Dynamic Optimization: Deterministic and Stochastic Models*. Springer International Publishing, 2017.

[HS00] W.J. Hopp and M.L. Spearman. *Factory Physics: Foundations of Manufacturing Management*. The McGraw-Hill/Irwin, 2 edition, 2000.

[HSH11] R. Haji, M. Saffari, and A. Haji. Inventory and Service System in a Two-Echelon Supply Chain Using Base Stock Policy. In *Proceedings of 22nd POMS Annual Conference Journal of Industrial and Systems Engineering*, Reno, Nevada, USA, paper number 020-0324, 2011.

[HWC75] U. Herzog, L. Woo, and K.M. Chandy. Solution of queueing problems by a recursive technique. *IBM Journal of Research and Development*, 19(3):295–300, May 1975.

[IS13] K.P.S. Isotupa and S.K. Samanta. A continuous review (s, Q) inventory system with priority customers and arbitrarily distributed lead times. *Mathematical and Computer Modelling*, 57:1259–1269, March 2013.

[Iso06] K.P.S. Isotupa. An (s, Q) Markovian inventory system with lost sales and two demand classes. *Mathematical and Computer Modelling*, 43:687–694, 2006.

[Iso07] K.P.S. Isotupa. Continuous Review (s, Q) Inventory System with Two Types of Customers. *International Journal of Agile Manufacturing*, 9(1):79–85, 2007.

[Iso11] K.P.S. Isotupa. An (s, Q) inventory system with two classes of customers. *International Journal of Operational Research*, 12:1–19, 2011.

[Iso15] K.P.S. Isotupa. Cost analysis of an $(S-1, S)$ inventory system with two demand classes and rationing. *Annals of Operations Research*, 233:411–421, January 2015.

[Jac57] J.R. Jackson. Networks of waiting lines. *Operations Research*, 5:518–521, 1957.

[Jac63] J.R. Jackson. Jobshop–like queueing systems. *Management Science*, 10:131–142, 1963.

[Jai68] N.K. Jaiswal. *Priority Queues*, volume 50 of *Mathematics in science and engineering*. Academic Press, New York – London, 1968.

[JCS16] L. Janssen, T. Claus, and J. Sauer. Literature review of deteriorating inventory models by key topics from 2012 to 2015. *International Journal of Production Economics*, 182:86–112, 2016.

[Jeg14] K. Jeganathan. Perishable inventory system at service facilities with multiple server vacations and impatient customers. *Journal of Statistics Applications & Probability Letters*, 3(3):63–73, 2014.

[Jeg15] K. Jeganathan. Linear retrial inventory system with second optional service under mixed priority service. *TWMS Journal of Applied and Engineering Mathematics*, 5(2):249–268, 2015.

[JKA16] K. Jeganathan, J. Kathiresan, and N. Anbazhagan. A retrial inventory system with priority customers and second optional service. *OPSEARCH*, 53:808–834, 2016.

[KA08] A. Krishnamoorthy and N. Anbazhagan. Perishable Inventory System at Service Facilities with N Policy. *Stochastic Analysis and Applications*, 26:120–135, 2008.

[Kar60] S. Karlin. Dynamic Inventory Policy with Varying Stochastic Demands. *Management Science*, 6(3):231–238, 1960.

[KD15] R. Krenzler and H. Daduna. Loss systems in a random environment: steady state analysis. *Queueing Systems*, 80(1-2):127–153, 2015.

[KDO16] R. Krenzler, H. Daduna, and S. Otten. Jackson networks in non-autonomous random environments. *Advances in Applied Probability*, 48(2), 2016.

[Kei79] J. Keilson. *Markov Chain Models – Rarity and Exponentiality.* Springer, New York, 1979.

[Kel76] F.P. Kelly. Networks of queues. *Advances in Applied Probability*, 8:416–432, 1976.

[Kel79] F.P. Kelly. *Reversibility and Stochastic Networks.* John Wiley and Sons, Chichester – New York – Brisbane – Toronto, 1979.

[Kle64] L. Kleinrock. *Communication nets – Stochastic message flow and delay.* McGraw–Hill, New York, 1964.

[Kle76] L. Kleinrock. *Queueing Theory*, volume II. John Wiley and Sons, New York, 1976.

[KLM11] A. Krishnamoorthy, B. Lakshmy, and R. Manikandan. A survey on inventory models with positive service time. *OPSEARCH*, 48(2):153–169, 2011.

[KMD15] A. Krishnamoorthy, R. Manikandan, and Shajin Dhanya. Analysis of a multiserver queueing-inventory system. *Advances in Operations Research*, Article ID 747328, 16 pages, 2015.

[KPC12] A. Krishnamoorthy, P.K. Pramod, and S.R. Chakravarthy. Queues with interruptions: a survey. *TOP*, 22:290–320, 2012.

[Kre16] R. Krenzler. *Queueing systems in a random environment.* PhD thesis, University of Hamburg, Department of Mathematics, 2016.

[KS16a] K. Karthikeyan and R. Sudhesh. Recent review article on queueing-inventory systems. *Research Journal of Pharmacy and Technology*, 9:2056–2066, 2016.

[KS16b] A. Krishnamoorthy and D. Shajin. Stochastic Decomposition in Retrial Queueing Inventory System. In *Proceedings of the 11th International Conference on Queueing Theory and Network Applications*, QTNA '16, pages 26:1–26:4, New York, USA, 2016. ACM.

[KSL16] A. Krishnamoorthy, D. Shajin, and B. Lakshmy. Product form solution for some queueing-inventory supply chain problem. *OPSEARCH*, pages 85–102, 2016.

[KSWD11] I.Z. Karaesmen, A. Scheller-Wolf, and B. Deniz. Managing perishable and aging inventories: Review and future research directions. In *Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook, Volume 1*, pages 393–436. Springer US, New York – Dordrecht – Heidelberg – London, 2011.

[Kum16] R. Kumar. Performance Analysis of a Markovian Queuing System with Reneging and Retention of Reneged Customers. In N. Anbazhagan, editor, *Stochastic Processes and Models in Operations Research*, Advances in Logistics, Operations, and Management Science, chapter 6, pages 97–105. IGI Global, 2016.

[KV11] A. Krishnamoorthy and C.N. Viswanath. Production inventory with service time and vacation to the server. *IMA Journal of Management Mathematics*, 22:33–45, January 2011.

[KV13] A. Krishnamoorthy and C.N. Viswanath. Stochastic decomposition in production inventory with service time. *European Journal of Operational Research*, 228:358–366, July 2013.

[KY14] F.P. Kelly and E. Yudovina. *Stochastic Networks.* Cambridge University Press, 2014.

[LC97] L. Liu and K.L. Cheung. Service constrained inventory models with random lifetimes and lead times. *Journal of the Operational Research Society*, 48:1022–1028, 1997.

[LFW14] M. Liu, M. Feng, and C.Y. Wong. Flexible service policies for a Markov inventory system with two demand classes. *International Journal of Production Economics*, 151:180–185, 2014.

[LR99] G. Latouche and V. Ramaswami, editors. *Introduction to Matrix Analytic Methods in Stochastic Modeling.* ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia, 1999.

[LXC13] J. Liu, F. Xi, and H. Chen. Control policies for a Markov queueing-inventory system with two demand classes. In *International Asia Conference on Industrial Engineering and Management Innovation (IEMI2012) Proceedings*, pages 1543–1550, 2013.

[LZ09] H. Li and Y.Q. Zhao. Exact tail asymptotics in a priority queue - characterizations of the preemptive model. *Queueing Systems*, 63:355–381, 2009.

[LZ12] M. Li and J. Zhang. Does inventory pooling improve customer service levels? *Operations Research Letters*, 40(2):96 – 98, 2012.

[Mel79] B. Melamed. Characterisation of Poisson streams in Jackson queueing networks. *Advances in Applied Probability*, 11:422–438, 1979.

[Mil58] R.G. Miller. Priority queues. Technical report, Stanford University, California, 1958.

[Min03] Stefan Minner. Multiple-supplier inventory models in supply chain management: A review. *International Journal of Production Economics*, 81:265–279, 2003. Proceedings of the Eleventh International Symposium on Inventories.

[Mor58] P.M. Morse. *Queues, Inventories and Maintenance.* Wiley, New York, 1958.

[MPB16] A.Z. Melikov, L.A. Ponomarenko, and S.A. Bagirova. Analysis of Queueing-Inventory Systems with Impatient Consume Customers. *Journal of Automation and Information Sciences*, 48(1):53–68, 2016.

[MS02] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks.* Wiley, Chichester, 2002.

[MSA07] P. Manuel, B. Sivakumar, and G. Arivarignan. A perishable inventory system with service facilities, MAP arrivals and PH-service times. *Journal of Systems Science and Systems Engineering*, 16:62–73, 2007.

[MSA08] P. Manuel, B. Sivakumar, and G. Arivarignan. A perishable inventory system with service facilities and retrial customers. *Computers and Industrial Engineering*, 54:484–501, 2008.

[Nah82] S. Nahmias. Perishable inventory theory: a review. *Operations Research*, 30(4):680–708, 1982.

[Nah10] S. Nahmias. Mathematical Models for Perishable Inventory Control. In J.J. Cochran, editor, *Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, 2010.

[Nah11] S. Nahmias. *Perishable inventory systems*. Springer, New York – Dordrecht – Heidelberg – London, 2011.

[Neu81] M.F. Neuts. *Matrix Geometric Solutions in Stochastic Models - An Algorithmic Approach*. John Hopkins University Press, Baltimore, MD, 1981.

[NL79] M.F. Neuts and D.M. Lucantoni. A Markovian queue with n servers subject to breakdowns and repairs. *Management Science*, 25(9):849–861, 1979.

[NR90] M.F. Neuts and B.M. Rao. Numerical investigation of a multiserver retrial model. *Queueing Systems*, 7(2):169–189, 1990.

[OKD14] S. Otten, R. Krenzler, and H. Daduna. Integrated models for production-inventory systems. `http://preprint.math.uni-hamburg.de/public/papers/prst/prst2014-01.pdf`, May 2014.

[OKD16] S. Otten, R. Krenzler, and H. Daduna. Models for integrated production-inventory systems: steady state and cost analysis. *International Journal of Production Research*, 54(20):6174–6191, 2016. (published online 2015).

[OKD17] S. Otten, R. Krenzler, and H. Daduna. Separable models for interconnected production-inventory systems. *submitted in Stochastic Models*, 2017.

[Ott13] S. Otten. Integrated models for queueing and inventory problems. Masterarbeit, Universität Hamburg, Department Mathematik, Schwerpunkt Mathematische Statistik und Stochastische Prozesse, January 2013.

[Pal89] M. Pal. The $(S-1, S)$ inventory model for deteriorating items with exponential lead time. *Calcutta Statistical Assocation Bulletin*, 38:149–150, 1989.

[Per90] H.G. Perros. Approximation algorithms for open queueing networks with blocking. In H. Takagi, editor, *Stochastic Analysis of Computer and Communication Systems*, pages 451–498. North-Holland, Amsterdam, 1990.

[Pib05] R. Pibernik. Advanced available-to-promies: Classification, selected methods and requirements for operations and inventory management. *International Journal of Production Economics*, pages 239–252, 2005.

[Pit79] B. Pittel. Closed exponential networks of queues with saturation: the Jackson-type stationary distribution and its asymptotic analysis. *Mathematics of Operations Research*, 4:357–378, 1979.

[PP98] D. Perry and M.J.M. Posner. An $(S-1, S)$ inventory system with fixed shelf life and constant lead times. *Operations Research*, 46(3):S65–S71, 1998.

[Raa91] F. Raafat. Survey of Literature on Continuously Deteriorating Inventory Models. *The Journal of the Operational Research Society*, 42(1):27–37, 1991.

[RBP13] R. Ravid, O.J. Boxma, and D. Perry. Repair systems with exchangeable items and the longest queue mechanism. *Queueing Systems*, pages 295–316, 2013.

[RM11] J.R. do Rego and M.A. de Mesquita. Spare parts inventory control: a literature review. *Produção*, 21(4):656–666, 2011.

[RW96] R. Rubio and L.M. Wein. Setting base stock levels using product-form queueing networks. *Operations Research*, 42:259–268, 1996.

[RZ17] J. Reed and B. Zhang. Managing capacity and inventory for multi-server make-to-stock queues. *Queueing Systems*, 86:61–94, 2017.

[SAH13] M. Saffari, S. Asmussen, and R. Haji. The M/M/1 queue with inventory, lost sale, and general lead times. *Queueing System*, 75:65–77, 2013.

[SBDK16] D. Shajin, B. Benny, T.G. Deepak, and A. Krishnamoorthy. A relook at queueing-inventory system with reservation, cancellation and common life time. *Communication in Applied Analysis*, 20:545–574, 2016.

[Sch73] R. Schassberger. *Warteschlangen*. Springer, Wien, 1973.

[Sch84] R. Schassberger. Decomposable stochastic networks: Some observations. In F. Baccelli and G. Fayolle, editors, *Modelling and Performance Evaluation Methodology*, volume 60 of *Lecture Notes in Control and Information Sciences*, chapter IV, pages 137–150. Springer, Berlin, 1984. Proceedings of the International Seminar, Paris, France, January 24-26, 1983.

[Sch02] C. Schneeweiß. *Einführung in die Produktionswirtschaft*. Springer–Verlag, Berlin – Heidelberg – New York, 2002.

[Sch04] M. Schwarz. *Stochastic Models in Inventory Theory*. Shaker Verlag, Aachen, 2004.

[SD03] C. Sauer and H. Daduna. Availability formulas and performance measures for separable degradable networks. *Economic Quality Control*, 18:165–194, 2003.

[SD04] C. Sauer and H. Daduna. Degradable networks with general up and down time distributions. In P. Buchholz, R. Lehnert, and M. Pioro, editors, *Proceedings of the 12th GI/ITG Conference on MMB together with 3rd PGTS*, pages 185–194, Berlin, 2004. VDE-Verlag.

[SD06] M. Schwarz and H. Daduna. Queueing systems with inventory management with random lead times and with backordering. *Mathematical Methods of Operations Research*, 64:383–414, 2006.

[Sen99] L.I. Sennott. *Stochastic dynamic programmnig and the control of queueing systems*. Wiley, New York, 1999.

[Ser99] R.F. Serfozo. *Introduction to Stochastic Networks*, volume 44 of *Applications of Mathematics*. Springer, New York, 1999.

[Ser13] B. Sericola. *Markov Chains: Theory and Applications*. John Wiley & Sons, London – Hoboken, 2013.

[SH09] M. Saffari and R. Haji. Queueing System with Inventory for Two-Echelon Supply Chain. *CIE International Conference*, pages 835–838, 2009.

[SHH11] M. Saffari, R. Haji, and F. Hassanzadeh. A queueing system with inventory and mixed exponentially distributed lead times. *The International Journal of Advanced Manufacturing Technology*, 53:1231–1237, 2011.

[Sin12] C. Singer. *Flexibilitätsmanagement zur Bewältigung von Unsicherheit in der Supply Chain*. Books on Demand, 2012.

[SKM10] H. Stadtler, C. Kilger, and H. Meyr. *Supply Chain Management und Advanced Planning*. Springer–Verlag, Heidelberg –Dordracht – London – New York, 2010.

[Slo11] S.A. Slotnick. Order acceptance and scheduling: A taxonomy and review. *European Journal of Operational Research*, 212:1–11, 2011.

[SN85] C.P. Schmidt and S. Nahmias. $(S-1,S)$ policies for perishable inventory. *Management Science*, 31(6):719–728, 1985.

[SPP98] E.A. Silver, D. Pyke, and R. Peterson. *Inventory Management and Production Planning and Scheduling*. John Wiley and Sons, Inc., Chichester – New York – Brisbane – Toronto – Singapore, 3 edition, 1998.

[SS00] N.H. Shah and Y.K. Shah. Literature survey on inventory models for deteriorating items. *Economics Annals*, 44:221–237, 2000.

[SSD+06] M. Schwarz, C. Sauer, H. Daduna, R. Kulik, and R. Szekli. $M/M/1$ queueing systems with inventory. *Queueing Systems: Theory and Applications*, 54:55–78, 2006.

[SSL92] K. Sigman and D. Simchi-Levi. Light traffic heuristic for an M/G/1 queue with limited inventory. *Annals of Operations Research*, 40:371–380, 1992.

[Tem05] H. Tempelmeier. *Bestandsmanagement in Supply Chains*. Books on Demand, 2005.

[Ter88] R.J. Tersine. *Principles of Inventory and Materials Management*. North-Holland, New York – Amsterdam – London, 3 edition, 1988.

[Tow75] D. Towsley. Local balance models of computer systems. Dissertation, University of Texas at Austin, Department of Computer Science, 1975.

[Tow80] D. Towsley. Queueing Network Models with State-Dependent Routing. *Journal of the Association for Computing Machinery*, 27(2):323–337, 1980.

[vdW89] J. van der Wal. Monotonicity of the throughput of a closed exponential queueing network in the number of jobs. *ORSpektrum*, 11:97–100, 1989.

[Ver04] M. Vernon. Quantitative system design. In P. Buchholz, R. Lehnert, and M. Pioro, editors, *Proceedings of the 12th GI/ITG Conference on MMB together with 3rd PGTS*, pages 11–13, Berlin, 2004. VDE-Verlag.

[VS06] P. Verhoef and L.M. Sloot. Out-of-stock: Reactions, antecedents, management solutions, and a future perspective. In M. Krafft and M.K. Mantrala, editors, *Retailing in the 21st Century: Current and Future Trends*, pages 239–253. Springer Publishers, Philadelphia, Pennsylvania, USA, 2006.

[Wal88] J. Walrand. *An introduction to queueing networks.* Prentice – Hall International Editions, Englewood Cliffs, 1988.

[WBSW15] J. Wang, O. Baron, and A. Scheller-Wolf. $M/M/c$ Queue with Two Priority Classes. *Operations Research*, 63(3):733–749, 2015.

[WLJ10] K. Wang, N. Li, and Z. Jiang. Queueing system with impatient customers: A review. In *Proceedings of 2010 IEEE International Conference on Service Operations and Logistics, and Informatics*, pages 82–87, July 2010.

[Woo94] M.E. Woodward. *Communication and Computer Networks: Modelling with Discrete–Time Queues.* IEEE Computer Society Press, Los Alamitos, Calif, 1994.

[YAJ15a] V.S.S. Yadavalli, N. Anbazhagan, and K. Jeganathan. A retrial inventory system with impatient customers. *Applied Mathematics & Information Sciences*, 9(2):637–650, 2015.

[YAJ15b] V.S.S. Yadavalli, N. Anbazhagan, and K. Jeganathan. A two heterogeneous servers perishable inventory system of a finite population with one unreliable server and repeated attempts. *Pakistan Journal of Statistics*, 31(1):135–158, 2015.

[YS09] H. Yang and L. Schrage. Conditions that cause risk pooling to increase inventory. *European Journal of Operational Research*, 192(3):837–851, 2009.

[Zaz94] M. Zazanis. Push and Pull Systems with External Demands. In *Proceedings of the 32nd Allerton Conference on Communication, Control, and Computing*, Allerton, Illinois, 1994.

[Zip00] P.H. Zipkin. *Foundations of inventory management.* Jeffrey J. Shelstad, 2000.

[ZL11] N. Zhao and Z. Lian. A queueing-inventory system with two classes of customers. *International Journal of Production Economics*, 129(1):225–231, 2011.

Abstract

Production processes are usually investigated using models and methods from queueing theory. Control of warehouses and their optimization rely on models and methods from inventory theory. Both theories are fields of Operations Research, but they comprise quite different methodologies and techniques. In classical Operations Research queueing and inventory theory are considered as disjoint research areas. On the other side, the emergence of complex supply chains ($\equiv$ production-inventory networks) calls for integrated production-inventory models as well as adapted techniques and evaluation tools.

Integrated production-inventory models are the focus of our research. We have developed Markovian stochastic models of several production-inventory systems, which are smooth enough to be amenable to solving the steady state problem explicitly with closed form expressions for the stationary distribution. Moreover, for most of the integrated production-inventory systems the obtained steady state is of so-called "product form", which reveals a certain decoupling of the components of the system for long time behaviour. This product form structure of the joint stationary distribution is often characterised as the global process being "separable". It is an important (but rather rare) property of complex systems. The simple structure of these steady states allows to apply "product form calculus", a widely used tool, which provides access to easy performance evaluation procedures.

Different from this standard product form equilibria in queueing networks, the steady state obtained for some integrated models is stratified. In the upper stratum, we obtain three vectors for production, inventory and supplier network. In the lower stratum, each of these vectors is composed of homogeneous coordinates. The product form inside the lower stratum resembles on one side (for the production subsystem) the independence structure of Jackson networks, and on the other side (for the inventory-replenishment subsystem) the conditional independence of Gordon-Newell networks. We briefly call this stratified product form "three-term product structure". On the other side, we mean by "two-term product form structure" that the steady states of the production network and the inventory-replenishment complex decouple asymptotically and the equilibrium for the production subsystem decomposes in true independent coordinates.

In the following, we will summarize our results in more detail for the investigated models of Part I and Part II. In Table E.1 on page 401, we compare the models from Part I to give an overview of the similarities and differences of the models, especially with regard to the modelling assumptions and the stationary distribution of the investigated models.

In Part I, we have developed Markov process models for several networks of production-inventory systems. More precisely, production-inventory systems at several locations are coupled by a common supplier. Demand of customers arrives at each production system

and is lost if the local inventory is depleted. To satisfy a customer's demand a server at the production system takes exactly one unit of raw material from the associated local inventory. The supplier manufactures raw material to replenish the local inventories, which are controlled by a continuous review base stock policy.

Chapter 2 to Chapter 4 focus on the research of the network's behaviour, where the supplier consists only of one machine (single server) and replenishes the inventories at all locations. The items of raw material are indistinguishable (exchangeable).

In Chapter 2, we have investigated the basic model. We have proven that the stationary distribution of the global integrated system is of two-term product form. Our model — with the send out procedure of the central supplier by a random selection scheme — can be seen as an approximation for the more complex model in Chapter 5, where the finished items are delivered exactly to the locations where the orders were generated and where the supplier consists of one workstation. The following conclusions can be drawn from our analytically obtained formulas: The global search for the vector of optimal base stock levels can be reduced to a set of independent optimization problems. In addition, the system exhibits various stochastic monotonicity properties. Additionally, for a subset of the parameter space, pooling is advisable. Moreover, ergodicity is determined by arrival and service intensities only. Finally, the system exhibits a strong insensitivity property for the inventory and replenishment process with respect to variation of the service rates.

Both last structural properties about ergodicity and insensitivity hold for the following models with product form results as well.

In Chapter 3, we have investigated a refined model, where routing of items depends on the on-hand inventory at the locations (with the aim to obtain "load balancing"). The systems under investigation differ with respect to the load balancing policy: In Section 3.3, we have considered strict priorities (i.e. the finished item of raw material is sent to the location(s) with the highest difference between the on-hand inventory and the capacity of the inventory) and in Section 3.4, we have considered weak priorities (i.e. the finished item of raw material is sent with greater probability to the location with higher difference between the on-hand inventory and the capacity of the inventory).

For the system with strict priorities for load balancing policy we have proven that the stationary distribution has a two-term product form of the marginal distributions of the production subsystem and the inventory-replenishment subsystem.

We have derived an explicit solution for the marginal distribution of the production subsystem. For the special case with base stock levels equal to one we have derived an explicit solution for the marginal distribution of the inventory-replenishment subsystem. For systems with two locations and base stock levels greater than one the marginal distribution of the inventory-replenishment subsystem can be obtained by a recursive algorithm.

For the system with weak priorities for load balancing policy we have derived the stationary distribution in explicit three-term product form.

In Chapter 4, we have analysed the basic model with perishable items, since in certain types of inventories, the items either perish, deteriorate or become obsolete.

In the first part of this chapter, we have studied single location production-inventory systems with perishable raw material, where the item of raw material that is in the production process cannot perish.

For phase-type distributed life time and queue-length-dependent arrival and service rates we have derived a sufficient and a necessary condition for ergodicity and have proven that the stationary distribution is not of product form. Furthermore, we have obtained structural properties of the stationary system which provide insights into the equilibrium behaviour of the system. An explicit expression of the complete stationary distribution is still an open problem. For the case with base stock level equal to one, we have obtained an explicit closed solution for the stationary distribution for the model with exponentially distributed life times and queue-length-independent arrival and service rates.

In the second part of this chapter, we have modified the system with exponentially distributed life times so that we have got product form results. The product form result is even true for a supply chain with $J > 1$ locations.

Furthermore, we have dealt with the question "Can we use the product form results from Section 4.3.2 to obtain simple product form bounds for the system with unknown non-product form stationary distribution in Section 4.2.3?". Our conjecture that we can find upper and lower product form bounds for the throughputs has been supported by our results for a system with base stock level $b = 1$. Under additional conditions for the system's parameters we can tackle even the case of $b \geq 2$.

The model can be considered as a special case of queueing systems in a random environment which we have introduced in Appendix D.1. We have derived a sufficient and a necessary condition for ergodicity and have proven that the stationary distribution is not of product form. Moreover, we have derived structural properties of the stationary system.

Chapter 5 to Chapter 7 is devoted to research of the network's behaviour of the more complex model where the finished items of raw material are delivered exactly to the locations where the orders were generated, i.e. they are not exchangeable.

In Chapter 5, we have investigated the basic model where the supplier is a complex network. We have proven that the stationary distribution of the global integrated system is of three-term product form. Furthermore, we have performed a cost analysis to find the optimal base stock levels.

In Chapter 6, we have aggregated the supplier network. We can substitute the complex supplier network only by one node — a supplier who consists of a symmetric server. The symmetric server enables to deal with non-exponential type-dependent service time distribution for different order types. For a complex supply chain with phase-type distributed service times at the supplier we have derived its stationary distributions of the joint queueing-inventory process in explicit three-term product form.

In Chapter 2 to Chapter 6, we have focused on base stock policies. In Chapter 7, we have investigated the (r, S)-policy. The systems under investigation differ with respect to the reorder level and the number of locations and workstations. In Section 7.3, we have analysed a system with $(0, S_j)$-policy with J locations and a supplier network consisting of M workstations and in Section 7.4 a system with $(1, S_j)$-policy with two locations and a supplier network consisting of one workstation. For these systems we have derived the stationary distributions of the joint queueing-inventory processes in explicit product form, which is of three-term product form for the $(0, S_j)$-policy and of two-term product form for the $(1, S_j)$-policy.

In Part II, we have studied a production-inventory system with two classes of customers and inventory management under lost sales with (r, Q)-policy in Chapter 10 and base stock policy in Chapter 11 where customers' arrivals are regulated by a flexible admission control. For each of these systems we have developed a Markov process description and have analysed the global balance equations and the existence of a stationary distribution.

We have derived structural properties of the steady state distribution which provide insights into the equilibrium behaviour of the systems but an explicit expression of the complete stationary distribution is still an open problem. For the system with base stock policy we have derived a sufficient condition for ergodicity by the Foster-Lyapunov stability criterion.

We have considered the special case of zero service time in Section 10.2 and in Section 11.3, which results in a pure inventory system and have determined the stationary distribution.

In Table E.1 on page 401, the symbols are used as follows.

Symbol	Meaning
π ✓	stationary distribution has an explicit solution
ξ ✓	only the marginal distribution of the production subsystem has an explicit solution
θ ✓	only the marginal distribution of the inventory-replenishment subsystem has an explicit solution
$(\star)$	no difference between exchangeable and location specific items since there is only one location

model chapter	number of locations	number of workstations	arrival stream (prod. syst.)		service time (prod. syst.)		replenishment policy		items of raw material						service time (workstation)			stationary distribution		
			Pois		exp				exchangeable?			perishable?			exp			product form?		
			queue-length-dependent?		queue-length-dependent?		base stock	(r, S)	yes		no	yes		no	queue-length-dependent?		PH	yes		no
			yes	no	yes	no			rand. routing	load balanc.		exp	PH		yes	no		two-term	three-term	
2	J	1		✓	✓		✓		✓					✓		✓		π ✓		
3	J	1		✓	✓		✓			✓				✓		✓		strict pol. $b = 1$: π ✓ $b > 1$: ξ ✓	weak pol. π ✓	
4.2.2	1	1		✓		✓	✓				$(\star)$		✓			✓				✓
	1	1	✓		✓		✓				$(\star)$		✓			✓				✓
4.2.3	1	1		✓		✓	✓				$(\star)$	✓				✓				✓
	1	1	✓		✓		✓				$(\star)$	✓				✓				✓
4.3	J	1		✓	✓		✓		✓			✓				✓		π ✓		
5	J	M		✓	✓		✓				✓			✓	✓				π ✓	
6.3	J	1		✓	✓		✓				✓			✓			✓ sym.		π ✓	
6.4	J	1		✓	✓		✓				✓			✓	✓ sym.				π ✓	
7.3	J	M		✓	✓			✓ $r_j = 0$			✓			✓	✓				π ✓	
7.4	2	1		✓	✓			✓			✓			✓		✓		π ✓		

Table E.1.: A summary of the modelling assumptions in the networks of production-inventory systems (Part I)

Zusammenfassung

Produktionsprozesse werden in der Regel mit Hilfe von Modellen und Methoden der Warteschlangentheorie untersucht, während bei der Kontrolle und Optimierung von Lagersystemen Modelle und Methoden der Lagerhaltungstheorie zum Einsatz kommen. Beide Theorien sind Gebiete des Operations Research, sie verwenden jedoch unterschiedliche Methoden und Techniken. Im klassischen Operations Research werden Warteschlangen- und Lagerhaltungstheorie als getrennte Forschungsgebiete betrachtet. Andererseits werden integrierte Produktions-Lagermodelle sowie geeignete Techniken und Bewertungswerkzeuge aufgrund des Entstehens komplexer Supply Chains ($\equiv$ Produktions-Lager-Netzwerke) benötigt.

In dieser Arbeit stehen integrierte Produktions-Lagermodelle im Fokus unserer Untersuchungen. Für verschiedene Produktions-Lager-Systeme haben wir stochastische Markov Modelle entwickelt, die es ermöglichen die stationäre Verteilung in geschlossener Form zu bestimmen. Für die meisten hier untersuchten integrierten Produktions-Lager-Systeme erhalten wir die stationäre Verteilung in sogenannter "Produktform". Diese zeigt bestimmte Entkopplungen von Komponenten des Systems für das Langzeitverhalten auf. Die Produktformstruktur der gemeinsamen stationären Verteilung wird oft dadurch charakterisiert, dass der globale Prozess separabel ist. Das ist eine wichtige (aber seltene) Eigenschaft komplexer Systeme. Die einfache Struktur dieser stationären Zustände ermöglicht es einen "Produktformkalkül" anzuwenden. Dies ist ein weitverbreitetes Werkzeug, das Zugang zu einfachen Leistungsbewertungsverfahren ermöglicht.

Im Unterschied zum Standard-Produktform-Gleichgewicht, ist der stationäre Zustand für einige unserer integrierten Modelle "geschichtet". In der oberen Schicht erhält man Vektoren für Produktion, Lager und Zulieferer-Netzwerk. In der unteren Schicht setzt sich jeder dieser Vektoren aus homogenen Koordinaten zusammen. Die Produktform innerhalb der unteren Schicht entspricht auf der einen Seite (für das Produktions-Teilsystem) der unabhängigen Struktur eines Jackson-Netzwerkes und auf der anderen Seite (für das Lager-Zulieferer-Teilsystem) der bedingten Unabhängigkeit eines Gordon-Newell-Netzwerkes. Wir nennen diese geschichtete Produktform "Drei-Term-Produktformstruktur". Hingegen sprechen wir von "Zwei-Term-Produktformstruktur", wenn sich die stationären Zustände des Produktions-Netzwerkes und des Lager-Zulieferer-Komplexes entkoppeln. Dabei zerfällt das Gleichgewicht für das Produktions-Teilsystem weiter in unabhängige Koordinaten.

Unsere Resultate der untersuchten Modelle aus Teil I und Teil II sollen im Folgenden zusammengefasst werden. Die Gemeinsamkeiten und Unterschiede der verschiedenen Modelle aus Teil I — insbesondere im Hinblick auf die Modellannahmen und die stationären Verteilungen — werden in Tabelle E.1 verdeutlicht.

In Teil I haben wir uns mit verschiedenen Netzwerken von Produktions-Lager-Systemen befasst, die wir als Markov-Prozesse modelliert haben. In diesen Netzwerken sind Produktions-Lager-Systeme an mehreren Standorten über einen Zulieferer miteinander verbunden. An jedem Produktionssystem kommen Kunden an. Falls das lokale Lager leer ist, gehen diese verloren. Damit der Bediener des Produktionssystems die Kundennachfrage erfüllen kann, benötigt er ein Teil aus dem angeschlossenen lokalen Rohmateriallager. Das Lager wird gemäß einer kontinuierlichen Base-Stock-Politik vom Zulieferer aufgefüllt, der das Rohmaterial herstellt.

In Kapitel 2 bis 4 haben wir das Verhalten dieser Netzwerke untersucht. Wir haben insbesondere angenommen, dass der Zulieferer nur aus einer Arbeitsstation (einem Bediener) besteht und dieser die Lager an allen Standorten auffüllt. Das Rohmaterial ist austauschbar, das bedeutet, dass die Auffüllungsprozedur nach einem zufälligen Auswahlsystem stattfinden kann.

Unsere Untersuchungen in Kapitel 2 wenden sich dem Grundmodell zu. Wir haben bewiesen, dass die stationäre Verteilung des gesamten Systems eine Zwei-Term-Produktformstruktur aufweist. Unser Modell — mit der zufallsgesteuerten Auffüllungsstrategie — kann als eine Approximation für das komplexere Modell in Kapitel 5 (mit einer Arbeitsstation) angesehen werden. In dem komplexeren Modell wird das Rohmaterial genau zu dem Standort geliefert, der die Bestellung aufgegeben hat. Aus unseren analytischen Ergebnissen erhalten wir die folgenden Resultate: (1) Die globale Suche nach dem Vektor der optimalen Base-Stock-Level kann auf eine Menge unabhängiger Optimierungsprobleme reduziert werden. (2) Das System weist mehrere stochastische Monotonie-Eigenschaften auf. (3) Pooling ist für eine Teilmenge des Parameterraums empfehlenswert. (4) Ergodizität wird nur durch die Ankunfts- und Bedienintensitäten bestimmt. (5) Das System weist starke Unempfindlichkeitseigenschaften für den Lager- und Auffüllprozess mit Variation der Bedienraten auf.

Die letzten beiden Struktureigenschaften (4) und (5) sind ebenfalls für die nachfolgenden Modelle mit Produktformstruktur gültig.

In Kapitel 3 haben wir das Grundmodell weiterentwickelt. Das Routing des Rohmaterials hängt nun von den Lagerbeständen an den verschiedenen Standorten ab. Ziel ist der "Belastungsausgleich durch Ausgleich der Lagerbestände". Hierzu haben wir zwei Lagerausgleichspolitiken analysiert: In Abschnitt 3.3 strikte Prioritäten und in Abschnitt 3.4 schwache Prioritäten. Strikte Prioritäten bedeutet, dass das produzierte Rohmaterial zu dem Standort gesendet wird, der die größte Differenz zwischen dem Lagerbestand und der Kapazität des Lagers hat. Hingegen drücken schwache Prioritäten aus, dass das produzierte Rohmaterial mit größerer Wahrscheinlichkeit zu dem Standort gesendet wird, der eine größere Differenz zwischen dem Lagerbestand und der Kapazität des Lagers hat.

Für das System mit strikten Prioritäten haben wir bewiesen, dass die stationäre Verteilung eine Zwei-Term-Produktformstruktur aufweist. Die Randverteilung des Produktions-Teilsystems ist explizit bekannt. Weiter ist die Randverteilung des Lager-Zulieferer-Teilsystems für den Spezialfall mit Base-Stock-Leveln gleich eins explizit gelöst. Für Systeme mit zwei Standorten und einem Base-Stock-Level größer als eins kann die Randverteilung des Lager-Zulieferer-Teilsystems durch einen rekursiven Algorithmus ermittelt werden.

Für das System mit schwachen Prioritäten haben wir die stationäre Verteilung in expliziter Drei-Term-Produktformstruktur ermittelt.

Häufig verderben oder altern Lagerbestände oder werden nicht mehr nachgefragt. Aus diesem Grund haben wir Kapitel 4 der Analyse des Grundmodells mit verderblichem Rohmaterial gewidmet.

Im ersten Teil des Kapitels haben wir ein Netzwerk mit einem einzigen Standort (Produktions-Lager-System) untersucht. Das Rohmaterial hat phasenverteilte Lebenszeiten. Es wird angenommen, dass das Rohmaterial, das sich im Produktionsprozess befindet, nicht verderben kann. Ferner werden warteschlangenlängenabhängige Ankunfts- und Bedienraten betrachtet.

Wir haben hinreichende und notwendige Ergodizitätsbedingungen hergeleitet und bewiesen, dass die stationäre Verteilung keine Produktformgestalt hat. Des Weiteren haben wir Struktureigenschaften für das System erhalten, die Einsichten in das Gleichgewichtsverhalten des Systems geben. Eine explizite Lösung für die stationäre Verteilung ist weiterhin ein offenes Problem. Für den Fall eines Base-Stock-Levels gleich eins konnten wir bereits eine explizite Lösung für das System mit exponentialverteilten Lebenszeiten und warteschlangenunabhängigen Ankunfts- und Bedienraten ermitteln.

Im zweiten Teil des Kapitels haben wir uns mit einer Modifikation des Systems mit exponentialverteilten Lebenszeiten befasst, dessen stationäre Verteilung Produktformgestalt besitzt. Die zugehörigen Resultate sind sogar für ein Netzwerk mit $J > 1$ Standorten gültig.

Abschließend haben wir uns mit der Frage beschäftigt, ob unsere Produktform-Resultate aus Abschnitt 4.3.2 genutzt werden können, um Produktformschranken für unser System aus Abschnitt 4.2.3 mit unbekannter Nicht-Produkformstruktur der stationären Verteilung zu erhalten. Unsere Vermutung, dass eine Monotonie-Eigenschaft der Durchsätze vorliegt, wird durch unsere Resultate für den Spezialfall mit Base-Stock-Level gleich eins sowie für Spezialfälle mit Base-Stock-Level größer gleich zwei unterstützt.

Die Modelle aus Kapitel 4 können als Spezialfall von Warteschlangensystemen in zufälliger Umwelt betrachtet werden, welche wir in Appendix D.1 eingeführt haben. Wir haben eine hinreichende und eine notwendige Bedingung für die Ergodizität hergeleitet. Zusätzlich haben wir bewiesen, dass die stationäre Verteilung keine Produktformgestalt hat. Ferner, haben wir Struktureigenschaften für das stationäre System ermittelt.

In Kapitel 5 bis 7 haben wir das Verhalten von komplexeren Netzwerken (Supply Chains) untersucht. Das vom Zulieferer produzierte Rohmaterial wird genau zu dem Standort gesendet, wo die Bestellung aufgegeben wurde. Somit ist in diesen Modellen das Rohmaterial nicht austauschbar.

In Kapitel 5 haben wir das Grundmodell untersucht, in dem der Zulieferer aus einem komplexen Netzwerk von Arbeitsstationen besteht. Wir haben bewiesen, dass die stationäre Verteilung des globalen integrierten Systems eine Drei-Term-Produkformstruktur aufweist. Ferner haben wir eine Kostenanalyse durchgeführt, mit dem Ziel die optimalen Base-Stock-Level zu bestimmen.

In Kapitel 6 haben wir das komplexe Zulieferernetzwerk aggregiert. Genauer gesagt, kann das komplexe Zulieferernetzwerk durch einen Knoten — einen Zulieferer, der aus einem symmetrischen Bediener besteht — substituiert werden. Der symmetrische Bediener ermöglicht es mit nicht-exponentiellen typabhängigen Bedienzeitverteilungen für verschiedene Typen zu arbeiten. Wir haben die stationäre Verteilung des gesamten Warteschlangen-Lager Prozesses einer komplexen Supply Chain mit phasenverteilten Bedienzeiten in expliziter Drei-Term-Produktformstruktur ermittelt.

In Kapitel 2 bis 6 haben wir uns mit Base-Stock-Politiken als Lagerhaltungspolitik beschäftigt. In Kapitel 7 haben wir uns der (r, S)-Politik als Lagerhaltungspolitik gewidmet. Die untersuchten Systeme unterscheiden sich hinsichtlich des Meldebestands, der Anzahl der Standorte sowie der Anzahl der Arbeitsstationen beim Zulieferer. In Abschnitt 7.3 haben wir die $(0, S_j)$-Politik mit J Standorten und einem Zulieferer-Netzwerk, das aus M Arbeitsstationen besteht, und in Abschnitt 7.4 die $(1, S_j)$-Politik mit zwei Standorten und einem Zulieferer, der aus einer Arbeitsstation besteht, behandelt. Für diese Systeme haben wir die stationären Verteilungen des Warteschlangen-Lager-Prozesses in expliziter Produktform bestimmt, welche für die $(0, S_j)$-Politik eine Drei-Term- und für die $(1, S_j)$-Politik eine Zwei-Term-Struktur aufweisen.

In Teil II haben wir Produktions-Lager-Systeme mit zwei Kundenklassen, Kundenverlust und (r, Q)-Politik in Kapitel 10 und Base-Stock-Politik in Kapitel 11 analysiert, in denen die Kundenankünfte durch eine flexible Zugangskontrolle reguliert werden. Diese Systeme haben wir als Markov-Prozesse modelliert und ihre globalen Gleichgewichtsgleichungen sowie die Existenz ihrer stationären Verteilung studiert.

Des Weiteren haben wir Struktureigenschaften für die Systeme gezeigt, die Einsichten in das Gleichgewichtsverhalten der Systeme geben. Allerdings sind explizite Lösungen der stationären Verteilungen noch ein offenes Problem. Für das System mit Base-Stock-Politik haben wir eine hinreichende Bedingung für Ergodizität mit Hilfe des Foster-Lyapunov Stabilitätskriteriums ermittelt.

In Abschnitt 10.2 und 11.3 haben wir reine Lagersysteme analysiert, die aus Setzen der Bedienzeiten gleich Null resultieren. Für diese Systeme haben wir die stationären Verteilungen bestimmt.

In Tabelle F.2, werden folgende Symbole verwendet.

Symbol	Bedeutung
π ✓	Stationäre Verteilung besitzt explizite Lösung
ξ ✓	Nur Randverteilung des Lager-Zulieferer-Teilsystems hat explizite Lösung
θ ✓	Nur Randverteilung des Lager-Zulieferer-Teilsystems hat explizite Lösung
$(\star)$	Keine Unterscheidung zwischen austauschbarem und standortspezifischem Rohmaterial, da das System nur einen Standort besitzt

Modell Kapitel	Anzahl der Stand-orte	Anzahl der Arbeits-stationen	Ankunftsstrom (Prod.syst.)		Bedienzeit (Prod.syst.)		Lagerhaltungs-politik		Rohmaterial						Bedienzeiten (Arbeitsstation)			stationäre Verteilung		
			Pois		exp		Base Stock	(r, S)	austauschbar?			verderblich?			exp		PH	Produktform?		
			Wschl'längen-abhängig?		Wschl'längen-abhängig?				ja		nein	ja		nein	Wschl'längen-abhängig?			ja		nein
			ja	nein	ja	nein			zufäll. Routing	Lager-ausgl.		exp	PH		ja	nein		zwei-Term	drei-Term	
2	J	1		✓	✓		✓		✓					✓		✓		π ✓		
3	J	1		✓	✓		✓			✓				✓		✓		strikte Pol. $b = 1$: π ✓ $b > 1$: ξ ✓	schw. Pol. π ✓	
4.2.2	1	1		✓		✓	✓				$(\star)$		✓			✓				✓
	1	1	✓		✓		✓				$(\star)$		✓			✓				✓
4.2.3	1	1		✓		✓	✓				$(\star)$	✓				✓				✓
	1	1	✓		✓		✓				$(\star)$	✓				✓				✓
4.3	J	1		✓	✓		✓		✓			✓				✓		π ✓		
5	J	M		✓	✓		✓				✓			✓	✓				π ✓	
6.3	J	1		✓	✓		✓				✓			✓			✓ sym.		π ✓	
6.4	J	1		✓	✓		✓				✓			✓	✓ sym.				π ✓	
7.3	J	M		✓	✓			✓ $r_j = 0$			✓			✓	✓				π ✓	
7.4	2	1		✓	✓			✓			✓			✓		✓		π ✓		

Tabelle F.2.: Zusammenfassung der Modellierungsannahmen in den Netzwerken von Produktions-Lager-Systemen (Part I)

List of publications derived from the dissertation up to now

Parts of this thesis are or will be published as following preprints and articles:

[OKD14] preprint about the basic production-inventory model with base stock policy and exchangeable items as well as the basic production-inventory model with base stock policy and location specific items. Parts are resp. will be published in [OKD16] and [OKD17].

[OKD16] journal article. It contains parts from [OKD14] focused on Section 3: *Reduced model with aggregated supplier network and indistinguishable orders.*

[OKD17] submitted article (under first revision). It contains parts from [OKD14] focused on Section 2: *Models for production-inventory systems: Multi-product systems.*