ORIGINAL RESEARCH

Load balancing in a network of queueing-inventory systems

Sonja Otten[1]

Abstract

We study a supply chain consisting of production-inventory systems at several locations which are coupled by a common supplier. Demand of customers arrives at each production system according to a Poisson process and is lost if the local inventory is depleted ("lost sales"). To satisfy a customer's demand a server at the production system needs 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. The routing of items depends on the on-hand inventory at the locations with the aim to obtain "load balancing". We show that the stationary distribution has a product form of the marginal distributions of the production subsystem and the inventory-replenishment subsystem. For the marginal distribution of the production subsystem we derive an explicit solution and for the marginal distribution of the inventory-replenishment subsystem we deduce an explicit solution or a recursive algorithm for some special cases.

Keywords Queueing networks · Inventory control · Base stock policy · Product-form steady state

Mathematics Subject Classification 60K25 · 68M20 · 90B22 · 90B05 · 90B06

1 Introduction

We consider a supply chain consisting of production-inventory systems at several locations which 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. We 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). The routing of items depends on the on-hand inventory at the locations with

✉ Sonja Otten
sonja.otten@tuhh.de

1 Institute of Mathematics, Hamburg University of Technology, Am Schwarzenberg-Campus 3, 21073 Hamburg, Germany

the aim to obtain "load balancing". More precisely, we consider 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.

Although we describe our system in terms of production and manufacturing, there are other applications where our model can be used, 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.

We use queueing theory and inventory theory to analyse the above-described production-inventory-replenishment system. Queueing theory and inventory theory are fields of Operations Research with different methodologies to optimise e.g. production processes and inventory control. In classical Operations Research, queueing theory and inventory theory are often considered as disjoint areas of research. On the other side, the emergence of complex supply chains calls for integrated queueing-inventory models, where, as usual, the production systems are modelled by queueing systems and the inventories are modelled by classical inventories with replenishment policies. These integrated queueing-inventory models are the focus of our present research. In particular, we develop a Markovian stochastic model of the production-inventory-replenishment system.

We are able to prove that the stationary distribution has a product form. This means 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. This product form structure of the joint stationary distribution is often characterised as the global process being "separable". Separability is an important (but rather rare) property of complex systems.

The paper is organised as follows. In Sect. 2, we describe the related literature. In Sect. 3, we introduce our integrated model for production and inventory management. To model our production-inventory-replenishment system as a queueing network, we make a number of simplifying assumptions. This enables us to compute the stationary distribution and to show in Sect. 4 that it is of product form. In Sect. 5 we show some structural properties of the stationary inventory-replenishment subsystem.

2 Related literature and own contributions

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 (1979) and Chao et al. (1999) for general networks of queues.

Special queueing networks, which model multi-station maintenance and repair systems, are investigated by Ravid et al. (2013) and Daduna (1990) and references therein. In these systems, circulating items are "exchangeable". This feature will occur in our model as well.

Literature on inventory theory is, similar to that on queueing theory, overwhelming, so we only point to some references closely related to our investigations. There are two extreme

cases of arriving customers' reactions in the situation that inventory is depleted when demand arrives [cf. Silver et al. (1998)]: 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 et al., 1998; Zipkin 2000, p. 40; Axsäter 2000). However, studies by Gruen et al. (2002) and Verhoef and Sloot (2006) 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 (2011). 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 Donselaar (2013) "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 (1996) and Zazanis (1994) 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 (2017) 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 (2005, 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" (Rego & Mesquita, 2011, p. 661).

Morse (1958, 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 Reed and Zhang (2017), etc.

Literature on integrated queueing-inventory models (i.e. queueing theory in combination with inventory theory) is also overwhelming. For a recent review we refer to Krishnamoorthy et al. (2021).

In this paper, we analyse an extension of the complex supply chain from Otten et al. (2016), where routing of items depends on the on-hand inventory at the locations with the aim to obtain "load balancing". It can be differed between two load balancing policies. In Otten (2017, Section 3.4) we consider 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. In this paper, we focus on 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. Therefore, we mention here some

Literature about load balancing policies with strict priorities: 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" (Daduna, 1987, p. 1) and

---

[1] The paper is from 1958.

"Unfortunately it turned out that introducing state-dependent routing into product form networks usually destroys the product-form of the steady state probabilities" (Daduna, 1987, p. 1). Product form solutions under state-dependent routing are, for example, found in Pittel (1979), Hordijk and Dijk (1984) and Schassberger (1984).

Several other stock allocation policies can, for example, be found in the article of Abouee-Mehrizi et al. (2014). 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 (1985, p. 624) and Towsley (1980, 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.

Chow and Kohler (1977, 1979) analyse the performance of two-processor distributed computer systems under several dynamic load balancing policies. 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 (Chow & Kohler, 1977) 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, which was first suggested by Herzog et al. (1975) for queueing problems. Their strategy "join the shorter queue without channel transfer" in Chow and Kohler's Model B (Chow & Kohler, 1977, pp. 42f.) is equivalent to our strict load balancing policy. Additionally, they allow a channel transfer in Model C (Chow & Kohler, 1977, pp. 42f.).

Flatto and McKean (1977) study also Chow and Kohler's Model B and derive by the generating function approach a complicated closed form solution (cf. Flatto & McKean, 1977, Section 3, p. 261), where the interarrival times of customers are exponentially distributed with rate 1.

Chow and Kohler (1979) present a generalization of the recursive solution technique. They apply the method to non-homogeneous (= heterogeneous) two-processor systems with special properties in Chow and Kohler (1979, Section IV, pp. 358f.) and present a sample system using the algorithm in Chow and Kohler (1979, Appendix, pp. 360f.). They mention that the generalization of recursive solution "(...) technique for three or more processors does not appear to be straightforward" (Chow & Kohler, 1979, p. 359).

Our main contributions are the following: We extend the network of queueing-inventory systems from Otten et al. (2016) by load balancing. 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 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 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. On the other hand, our work is an extension of the investigations of Chow and Kohler (1977, 1979): 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 (1979, Section IV, pp. 358f.) (for more details see Remark 6). Therefore, we construct a new algorithm.

Notations and conventions:

- $\mathbb{N} := \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$.
- 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". For a set A we denote by $|A|$ the number of elements in A.
- $1_{\{expression\}}$ is the indicator function which is 1 if *expression* is true and 0 otherwise.
- Empty sums are 0, and empty products are 1.
- Throughout this article 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 ($=$ 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).

3 Description of the general model

The supply chain of interest is depicted in Fig. 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).

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 distributed with mean 1. 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 $\mathbf{b} := (b_j : j \in \overline{J})$.

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
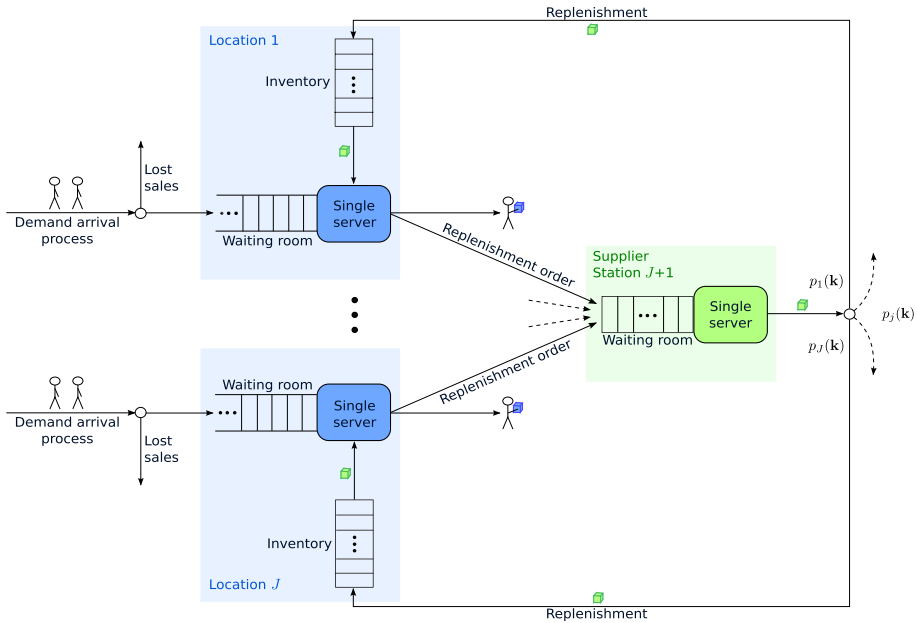
Fig. 1 Supply chain with base stock policies and load balancing policies

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, to location j, $j \in \overline{J}$, if the state of the inventory-replenishment subsystem is $\mathbf{k} := (k_1, \ldots, k_J, k_{J+1})$, i.e. k_j is the size of the inventory at location $j \in \overline{J}$ and k_{J+1} the number of replenishment orders at the central supplier. We consider the following *load balancing policy* with strict priorities: The finished item of raw material is sent to location $j \in \overline{J}$ with probability

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

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.

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

$$K := \left\{ (k_1, \ldots, k_J, k_{J+1}) | 0 \leq k_j \leq b_j, \ j = 1, \ldots, J, \ k_{J+1} = \sum_{j=1}^{J} (b_j - k_j) \right\}$$

$$\subset \mathbb{N}_0^{J+1}.$$

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.

Discussion of the modelling assumptions We have imposed several simplifying assumptions on the production-inventory-replenishment 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-replenishment system.

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

4 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 $(\mathbf{n}, \mathbf{k}) \in E$:

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

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

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

Proposition 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 $\mathbf{x} := (x(\mathbf{n}, \mathbf{k}) : (\mathbf{n}, \mathbf{k}) \in E)$ with*

$$x(\mathbf{n}, \mathbf{k}) = \widetilde{\xi}(\mathbf{n}) \cdot \widetilde{\theta}(\mathbf{k}), \tag{1}$$

where

$$\widetilde{\xi}(\mathbf{n}) = \prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j), \qquad \widetilde{\xi}_j(n_j) = \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0, \; j \in \overline{J}, \tag{2}$$

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

$$x(\mathbf{n}, \mathbf{k}) \cdot \left(\sum_{i \in \overline{J}} \lambda_i \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}} + \sum_{i \in \overline{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i < b_i\}} \right)$$

$$= \sum_{i \in \overline{J}} x(\mathbf{n} - \mathbf{e}_i, \mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}}$$

$$+ \sum_{i \in \overline{J}} x(\mathbf{n} + \mathbf{e}_i, \mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \mu_i(n_i + 1) \cdot 1_{\{k_i < b_i\}}$$

$$+ \sum_{i \in \overline{J}} x(\mathbf{n}, \mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot 1_{\{k_i > 0\}}.$$

It has to be shown that the stationary measure from Proposition 1 satisfies these global balance equations.

Substitution of (1) and (2) into the global balance equations directly leads to

$$\left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\theta}(\mathbf{k})$$

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

$$= \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J} \setminus \{i\}} \widetilde{\xi}_j(n_j) \right) \widetilde{\xi}_i(n_i - 1) \cdot \widetilde{\theta}(\mathbf{k}) \cdot \lambda_i \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}}$$

$$+ \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J} \setminus \{i\}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\xi}_i(n_i + 1) \cdot \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \mu_i(n_i + 1) \cdot 1_{\{k_i < b_i\}}$$

$$+ \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i \cdot 1_{\{k_i > 0\}}.$$

By substitution of (2) we obtain

$$
\left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\theta}(\mathbf{k})
$$

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

$$
= \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\theta}(\mathbf{k}) \cdot \mu_i(n_i) \cdot 1_{\{n_i > 0\}} \cdot 1_{\{k_i > 0\}}
$$

$$
+ \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i \cdot 1_{\{k_i < b_i\}}
$$

$$
+ \sum_{i \in \overline{J}} \left(\prod_{j \in \overline{J}} \widetilde{\xi}_j(n_j) \right) \cdot \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i \cdot 1_{\{k_i > 0\}}.
$$

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

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

$$
= \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i \cdot 1_{\{k_i < b_i\}}
$$

$$
+ \sum_{i \in \overline{J}} \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot 1_{\{k_i > 0\}}. \tag{3}
$$

An inspection of the system (3) 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 $\mathbf{k} \in K$:

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

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

The Markov process generated by $\mathbf{Q}_{red}$ is irreducible on K and therefore (3) has a solution which is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. $\qquad\square$

Remark 1 In Sect. 4.1.1, 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 Sect. 4.1.2.

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

Definition 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

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

Theorem 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

$$\pi(\mathbf{n}, \mathbf{k}) = \xi(\mathbf{n}) \cdot \theta(\mathbf{k}), \tag{4}$$

with

$$\xi(\mathbf{n}) = \prod_{j \in \overline{J}} \xi_j(n_j), \qquad \xi_j(n_j) = C_j^{-1} \prod_{\ell=1}^{n_j} \frac{\lambda_j}{\mu_j(\ell)}, \quad n_j \in \mathbb{N}_0, \ j \in \overline{J}, \tag{5}$$

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).

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 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 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 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 1. □

Remark 2 The expression (4) 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 (5) 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 (cf. Jackson (1957)) 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 (4) has been unexpected to us.

Our production-inventory-replenishment system can be considered as a "Jackson network in a random environment" in Krenzler et al. (2016, Section 4). We can interpret the inventory-replenishment subsystem, which contributes via θ to Theorem 2, as a "random environment" for the production network of nodes $\overline{J}$, which is a Jackson network of parallel servers. Taking into account the results of Krenzler et al. (2016, 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), is a consequence of that Theorem 4.1.

Our direct proof of Proposition 1 is much shorter than embedding the present model into the general framework of Krenzler et al. (2016).

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.

Ergodicity: As shown in Theorem 2, 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.

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 the estimation errors in determining the $\mu_j(\cdot)$. Vice versa, the distribution $\xi_j(n_j)$ is robust against changes in the inventory-replenishment network as long as the demand intensity and the service rates are maintained.

Remark 3 A cost analysis can be performed as for the model without load balancing in Otten et al. (2016, pp. 11f.). Note a corrected version of the definition of the cost function can be found in Otten (2017, Section 2.5, pp. 25f.).

4.1 Calculation of $\widetilde{\theta}$

4.1.1 Special case: $b_j = 1, j \in \bar{J}$

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 $b_j = 1$, $j \in \bar{J}$. In this special case, our model with strict priorities is identical to the model with weak priorities in Otten (2017, Section 3.4), where the finished item of raw material is sent with a greater probability to the location with higher difference between the on-hand inventory and the capacity of the inventory. For the sake of completeness we present the proof here.

We recall the notation for the inventory-replenishment subsystem

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

Furthermore, in this special case it holds $p_i(\mathbf{k}) = \frac{1}{J - \sum_{j \in \bar{J}} k_j}$ if $k_i = 0, i \in \bar{J}$.

Proposition 3 *The strictly positive measure* $\widetilde{\theta} = (\widetilde{\theta}(\mathbf{k}) : \mathbf{k} \in K)$ *of the inventory-replenishment subsystem with base stock levels* $b_1 = \cdots = b_J = 1$ *is given by*

$$\widetilde{\theta}(\mathbf{k}) = \widetilde{\theta}(k_1, k_2, \ldots, k_J, k_{J+1}) = \left(\prod_{\ell=0}^{(\sum_{j \in \bar{J}} k_j) - 1} \frac{1}{J - \ell} \right) \cdot \left(\prod_{j \in \bar{J}} \left(\frac{1}{\lambda_j} \right)^{k_j} \right) \cdot \left(\frac{1}{\nu} \right)^{k_{J+1}}$$

(6)

Remark 4 As we have mentioned before (3) has a solution which is unique up to a multiplicative constant, which yields $\widetilde{\theta}$. In Otten (2017, Proposition 3.3.6, p. 50) another stationary measure $\widetilde{\theta}_1$ is presented which differs from our stationary measure up to a multiplicative constant, i.e. $\widetilde{\theta} = \left(\frac{1}{\nu} \right)^J \cdot \widetilde{\theta}_1$.

Proof It has to be shown that the stationary measure (6) satisfies the global balance equations $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$, which are given for $\mathbf{k} \in K$ by

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

$$= \sum_{i \in \bar{J}} \widetilde{\theta}(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}) \cdot \lambda_i \cdot 1_{\{k_i=0\}}$$

$$+ \sum_{i \in \bar{J}} \widetilde{\theta}(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot \nu \cdot p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}) \cdot 1_{\{k_i=1\}}$$

$$\Leftrightarrow \widetilde{\theta}(\mathbf{k}) \cdot \left(\sum_{i \in \bar{J}} \lambda_i \cdot 1_{\{k_i=1\}} + \sum_{i \in \bar{J}} \nu \cdot p_i(\mathbf{k}) \cdot 1_{\{k_i=0\}} \right)$$

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

$$\cdot \underbrace{\frac{1}{J - \sum_{j \in \bar{J}} k_j}}_{=p_i(\mathbf{k})} \cdot \frac{1}{\lambda_i} \cdot \nu \cdot \lambda_i \cdot 1_{\{k_i=0\}}$$

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

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

The right hand-side of the last equation is

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

□

Remark 5 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 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 $\{1, \ldots, J\}$ it 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) reveal directly that this symmetry property holds in this case as well.

4.1.2 Special case: two locations and $b_j > 1, j \in \overline{J}$

In this section, we assume that there are two heterogeneous locations with base stock levels $b_1 \geq b_2$, where $b_1, b_2 > 1$ and arrival rates $\lambda_1, \lambda_2, \nu > 0$. The state transition diagram for such a system is presented in Fig. 2.

To obtain $\widetilde{\theta}(\mathbf{k})$, $\mathbf{k} \in K$, from the global balance equations $\widetilde{\theta} \cdot \mathbf{Q}_{red} = \mathbf{0}$ (cf. Eq. (3)) we use the recursive solution technique. This technique uses the fact that the steady state probabilities of the system can sometimes be expressed in terms of other steady state probabilities. Consequently, this leads to a reduction of the number of unknowns in the global balance equations. For our system the recursive method is described by the following algorithm.

κ is a variable which represents a temporarily unknown probability and GBE is used to denote a global balance equation. 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:

$$\widetilde{\theta}\left(\overbrace{k_1, \; k_2}^{\substack{\text{inventories} \\ \text{at locations}}} \right) := \widetilde{\theta}\left(\overbrace{k_1, \; k_2}^{\substack{\text{inventories} \\ \text{at locations}}} \overbrace{k_3}^{\text{supplier}} \right)$$
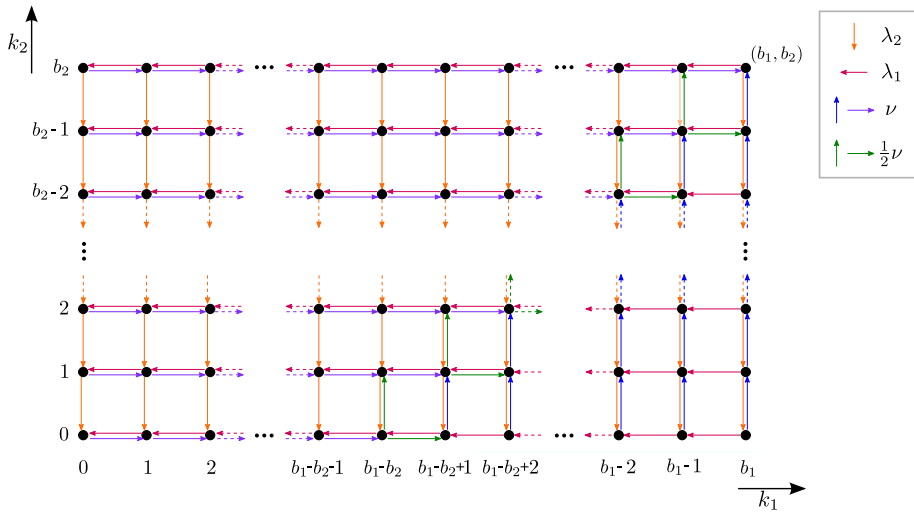
Fig. 2 State transition diagram of a system with two heterogeneous locations

and hence,

$$p_i(\mathbf{k} - \mathbf{e}_i) := p_i(\mathbf{k} - \mathbf{e}_i + \mathbf{e}_{J+1}), \qquad p_i(\mathbf{k} + \mathbf{e}_i) := p_i(\mathbf{k} + \mathbf{e}_i - \mathbf{e}_{J+1}), \ i \in \overline{J}.$$

A few steps of the algorithm are visualised in the state transition diagram in Fig. 3 and the algorithm is illustrated by an example in Appendix A, which is motivated by the computational example in Chow and Kohler (1979, Appendix, pp. 360f.). Furthermore, a detailed explanation of the algorithm can be found in Otten (2017, Appendix C.1, pp. 275–311).

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 $k_1 = 0, \ldots, b_1 - b_2 + 1$
      (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 $k_1 = b_1 - b_2, \ldots, 1$ (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

$$\widetilde{\theta}(k_1, k_2) \leftarrow \frac{\widetilde{\theta}(k_1, k_2)}{\sum_{k_1=0}^{b_1} \sum_{k_2=0}^{b_2} \widetilde{\theta}(k_1, k_2)}$$

Remark 6 Chow and Kohler apply a recursive solution technique to non-homogeneous (= heterogeneous) twoprocessor systems with special properties in (Chow & Kohler, 1979, Section IV, pp. 358f.). One special property of the load balancing policy that enables to use their 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 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 Fig. 2.

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 Fig. 4.
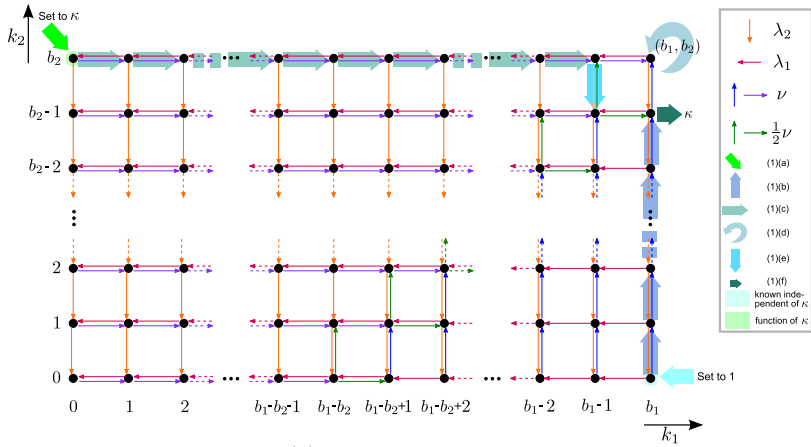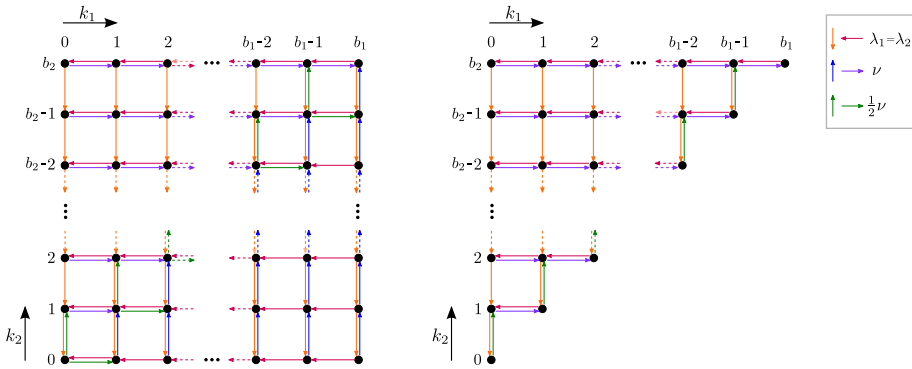
(a) First loop with $k_2 = b_2$
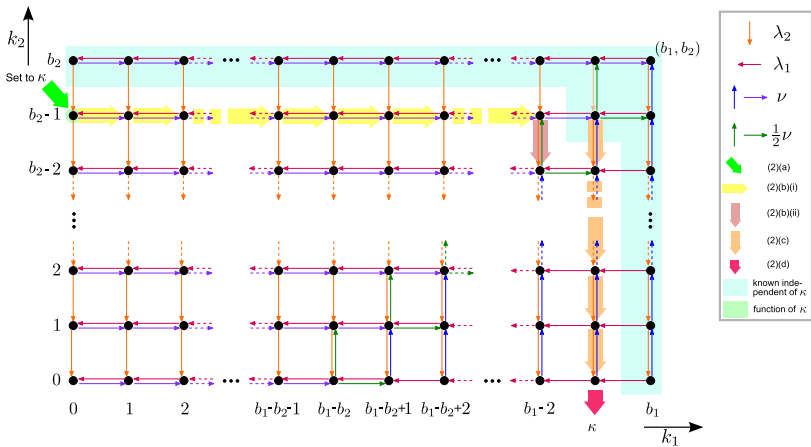
(b) Second loop with $k_2 = b_2 - 1$
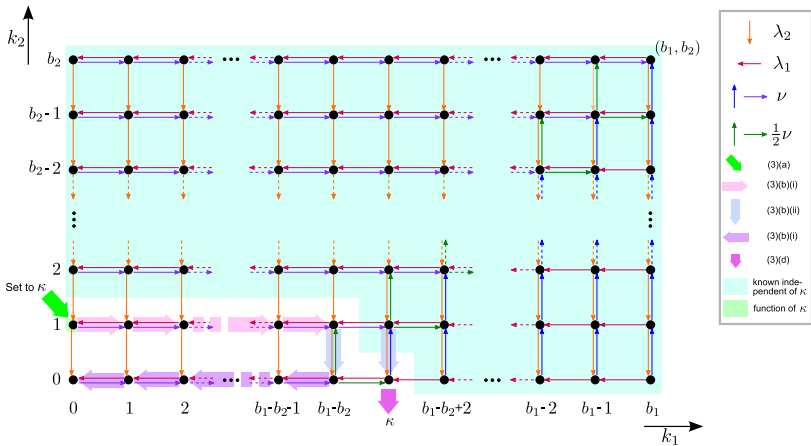
(c) Last loop with $k_2 = 1$

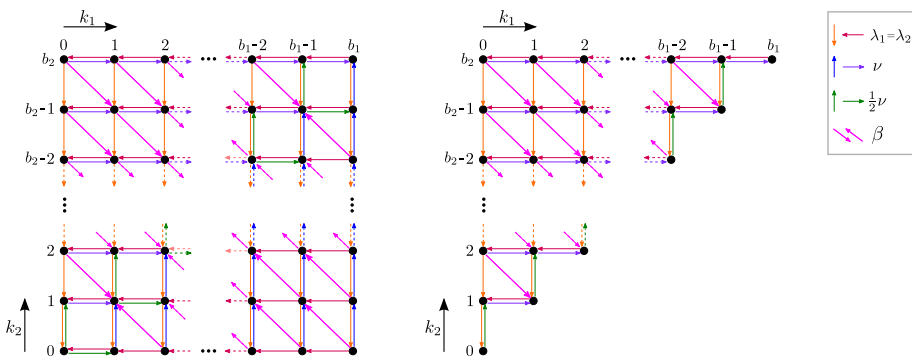**Fig. 3** Visualisation of a few steps of the algorithm

Fig. 4 State transition diagram (left) and folded state transistion diagram (right) of a system with two homogeneous locations

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 Fig. 4. Consequently, it holds $\widetilde{\theta}(k_1, k_2) = \widetilde{\theta}(k_2, k_1)$.

We can 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 ($=$ 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 Fig. 5.

Fig. 5 State transition diagram (left) and folded state transistion diagram (right) of a system with two homogeneous locations and transfer channel

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

Proposition 4 *For the inventory process holds*

$$P(Y_1 = \ell) = P(Y_j = \ell), \quad j \in \overline{J} \setminus \{1\}, \ \ell \in \{0, \ldots, b_1\},$$
$$P(Y_1 = \ell) \cdot \lambda_1$$
$$= P(Y_j = \ell - 1 \text{ for } j \in \overline{J}) \cdot \frac{\nu}{J}$$
$$+ \sum_{i=1}^{J-1} P(Y_j = \ell - 1 \text{ for } j = 1, \ldots, i \text{ and } \ell - 1 < Y_k \leq b_k \text{ for } k = i+1, \ldots, J)$$
$$\cdot \binom{J-1}{i-1} \cdot \frac{\nu}{i}, \qquad \ell \in \{1, \ldots, b_1\}. \tag{7}$$

Proof It holds

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

because of the symmetry property for homogeneous locations (see Remark 5) and because

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

The Eq. (7) can be proven by the cut-criterion for positive recurrent processes (see Kelly, 1979, Lemma 1.4, p. 8).

For $\ell \in \{1, \ldots, b_1\}$, 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

$$\left\{ (k_1, k_2, \ldots, k_J) : k_1 \in \{0, 1, \ldots, \ell - 1\}, \ k_j \in \{0, \ldots, b_j\}, \ j \in \{2, \ldots, J\} \right\},$$
$$\left\{ (\widetilde{k}_1, \widetilde{k}_2, \ldots, \widetilde{k}_J) : \widetilde{k}_1 \in \{\ell, \ldots, b_1\}, \ \widetilde{k}_j \in \{0, \ldots, b_j\}, \ j \in \{2, \ldots, J\} \right\}, \ \ell \in \{1, \ldots, b_1\}.$$

---

[2] It should be noted that $\theta(k_1, k_2, \ldots, k_J, k_{J+1}) = P(Y_1 = k_1, Y_2 = k_2, \ldots, Y_J = k_J, W_{J+1} = k_{J+1}) = P(Y_1 = k_1, Y_2 = k_2, \ldots, Y_J = k_J)$ 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 $\ell \in \{1, \ldots, b_1\}$

$$
\sum_{k_1=0}^{\ell-1} \sum_{k_2=0}^{b_2} \cdots \sum_{k_J=0}^{b_J} \sum_{\widetilde{k}_1=\ell}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \cdots \sum_{\widetilde{k}_J=0}^{b_J} \theta(k_1, k_2, \ldots, k_J)
$$

$$
\cdot q_{red}\left((k_1, k_2, \ldots, k_J); (\widetilde{k}_1, \widetilde{k}_2, \ldots, \widetilde{k}_J)\right)
$$

$$
= \sum_{\widetilde{k}_1=\ell}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \cdots \sum_{\widetilde{k}_J=0}^{b_J} \sum_{k_1=0}^{\ell-1} \sum_{k_2=0}^{b_2} \cdots \sum_{k_J=0}^{b_J} \theta(\widetilde{k}_1, \widetilde{k}_2, \ldots, \widetilde{k}_J)
$$

$$
\cdot q_{red}\left((\widetilde{k}_1, \widetilde{k}_2, \ldots, \widetilde{k}_J); (k_1, k_2, \ldots, k_J)\right)
$$

$$
\Leftrightarrow \quad \theta(\ell-1, \ell-1, \ldots, \ell-1) \cdot \frac{\nu}{J} \tag{8}
$$

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

$$
= \sum_{\widetilde{k}_2=0}^{b_2} \cdots \sum_{\widetilde{k}_J=0}^{b_J} \theta(\ell, \widetilde{k}_2, \ldots, \widetilde{k}_J) \cdot \lambda_1. \tag{9}
$$

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 $\left\{(\ell-1, k_2, \ldots, k_J) : k_j \in \{\ell-1, \ldots, b_j\}, \ j \in \{2, \ldots, J\}\right\}$, to $\left\{(\ell, \widetilde{k}_2, \ldots, \widetilde{k}_J) : \widetilde{k}_j \in \{\ell-1, \ldots, b_j\}, \ j \in \{2, \ldots, J\}\right\}$, $\ell \in \{1, \ldots, b_1\}$.

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 Eq. (8) all locations have exactly $\ell - 1$ items in the inventory and in Eq. (9) i states how many locations have exactly $\ell - 1$ items in the inventory. This results in the factor $\frac{1}{i}$, 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 $\ell \in \{1, \ldots, b_1\}$

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

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

$$
+ \sum_{i=1}^{J-1} P(Y_j = \ell - 1 \text{ for } j = 1, \ldots, i \text{ and } \ell - 1 < Y_k \leq b_k \text{ for } k = i+1, \ldots, J)
$$

$$
\cdot \binom{J-1}{i-1} \cdot \frac{\nu}{i}.
$$

$\square$

Heterogeneous locations

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 7 From Eq. (10) in the following proposition follows

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

Proposition 5 *For the inventory process holds for* $\ell_1 = 1, \ldots, b_1 - b_2$

$$P(Y_1 = \ell_1) \cdot \lambda_1 = P(Y_1 = \ell_1 - 1) \cdot \nu, \tag{10}$$

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

$$P(Y_1 = \ell_1) \cdot \lambda_1 = P(Y_1 = \ell_1 - 1, Y_2 = \ell_1 - 1 - b_1 + b_2) \cdot \frac{1}{2}\nu$$
$$+ P(Y_1 = \ell_1 - 1, Y_2 > \ell_1 - 1 - b_1 + b_2) \cdot \nu, \tag{11}$$

for $\ell_1 = b_1$

$$P(Y_1 = b_1) \cdot \lambda_1 = P(Y_1 = b_1 - 1, Y_2 = b_2 - 1) \cdot \frac{1}{2}\nu$$
$$+ P(Y_1 = b_1 - 1, Y_2 = b_2) \cdot \nu, \tag{12}$$

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

$$P(Y_2 = \ell_2) \cdot \lambda_2 = P(Y_1 = b_1 - b_2 + \ell_2 - 1, Y_2 = \ell_2 - 1) \cdot \frac{1}{2}\nu$$
$$+ P(Y_1 > b_1 - b_2 + \ell_2 - 1, Y_2 = \ell_2 - 1) \cdot \nu. \tag{13}$$

Proof The equations can be proven by the cut-criterion for positive recurrent processes (see Kelly, 1979, Lemma 1.4, p. 8).

For $\ell_1 \in \{1, \ldots, b_1 - b_2\}$, Eq. (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

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

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

$$\sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \theta(k_1, k_2) \cdot q_{red}\left((k_1, k_2); (\widetilde{k}_1, \widetilde{k}_2)\right)$$

$$= \sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \theta(\widetilde{k}_1, \widetilde{k}_2) \cdot q_{red}\left((\widetilde{k}_1, \widetilde{k}_2); (k_1, k_2)\right)$$

$$\Leftrightarrow \underbrace{\sum_{k_2=0}^{b_2} \theta(\ell_1 - 1, k_2) \cdot \nu}_{=P(Y_1=\ell_1-1)} = \underbrace{\sum_{\widetilde{k}_2=0}^{b_2} \theta(\ell_1, \widetilde{k}_2) \cdot \lambda_1}_{=P(Y_1=\ell_1)}.$$

---

[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 $k_3 = (b_1 + b_2) - (k_1 + k_2)$.

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 $\ell_1 \in \{b_1 - b_2 + 1, \ldots, b_1 - 1\}$, Eq. (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 $\ell_1 - 1$ or greater than $\ell_1 - 1$, i.e. into the sets

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

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

$$\sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \theta(k_1, k_2) \cdot q_{red}\left((k_1, k_2); (\widetilde{k}_1, \widetilde{k}_2)\right)$$

$$= \sum_{\widetilde{k}_1=\ell_1}^{b_1} \sum_{\widetilde{k}_2=0}^{b_2} \sum_{k_1=0}^{\ell_1-1} \sum_{k_2=0}^{b_2} \theta(\widetilde{k}_1, \widetilde{k}_2) \cdot q_{red}\left((\widetilde{k}_1, \widetilde{k}_2); (k_1, k_2)\right)$$

$$\Leftrightarrow \quad \underbrace{\sum_{k_2=b_2-(b_1-\ell_1)}^{b_2} \theta(\ell_1 - 1, k_2) \cdot \ \nu}_{=P(Y_1=\ell_1-1, Y_2 > b_2-(b_1-(\ell_1-1)))} + \underbrace{\theta(\ell_1 - 1, b_2 - (b_1 - (\ell_1 - 1)))}_{=P(Y_1=\ell_1-1, Y_2=b_2-(b_1-(\ell_1-1)))} \cdot \frac{1}{2}\nu$$

$$= \underbrace{\sum_{\widetilde{k}_2=0}^{b_2} \theta(\ell_1, \widetilde{k}_2) \cdot \lambda_1}_{=P(Y_1=\ell_1)}.$$

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

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

For $\ell_1 = b_1$, Eq. (12) 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

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

Then, it follows for

$$\sum_{k_1=0}^{b_1-1}\sum_{k_2=0}^{b_2}\sum_{\tilde{k}_1=b_1}^{b_1}\sum_{\tilde{k}_2=0}^{b_2}\theta(k_1,k_2)\cdot q_{red}\left((k_1,k_2);(\tilde{k}_1,\tilde{k}_2)\right)$$

$$=\sum_{\tilde{k}_1=b_1}^{b_1}\sum_{\tilde{k}_2=0}^{b_2}\sum_{k_1=0}^{b_1-1}\sum_{k_2=0}^{b_2}\theta(\tilde{k}_1,\tilde{k}_2)\cdot q_{red}\left((\tilde{k}_1,\tilde{k}_2);(k_1,k_2)\right)$$

$$\Leftrightarrow \quad \underbrace{\theta(b_1-1,b_2)}_{=P(Y_1=b_1-1,Y_2=b_2)}\cdot v + \underbrace{\theta(b_1-1,b_2-1)}_{=P(Y_1=b_1-1,Y_2=b_2-1)}\cdot\frac{1}{2}v = \underbrace{\sum_{\tilde{k}_2=0}^{b_2}\theta(b_1,\tilde{k}_2)}_{=P(Y_1=b_1)}\cdot\lambda_1.$$

Hence, we have shown

$$P(Y_1=b_1)\cdot\lambda_1 = P(Y_1=b_1-1,Y_2=b_2-1)\cdot\frac{1}{2}v + P(Y_1=b_1-1,Y_2=b_2)\cdot v.$$

For $\ell_2\in\{1,\dots,b_2\}$, Eq. (13) 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 ℓ_2-1 or greater than ℓ_2-1, i.e. into the sets

$$\left\{(k_1,k_2):k_1\in\{0,\dots,b_1\},\ k_2\in\{0,1,\dots,\ell_2-1\}\right\},$$
$$\left\{(\tilde{k}_1,\tilde{k}_2):\tilde{k}_1\in\{0,\dots,b_1\},\ \tilde{k}_2\in\{\ell_2,\dots,b_2\}\right\},\quad \ell_2\in\{1,\dots,b_2\}.$$

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

$$\sum_{k_1=0}^{b_1}\sum_{k_2=0}^{\ell_2-1}\sum_{\tilde{k}_1=0}^{b_1}\sum_{\tilde{k}_2=\ell_2}^{b_2}\theta(k_1,k_2)\cdot q_{red}\left((k_1,k_2);(\tilde{k}_1,\tilde{k}_2)\right)$$

$$=\sum_{\tilde{k}_1=0}^{b_1}\sum_{\tilde{k}_2=\ell_2}^{b_2}\sum_{k_1=0}^{b_1}\sum_{k_2=0}^{\ell_2-1}\theta(\tilde{k}_1,\tilde{k}_2)\cdot q_{red}\left((\tilde{k}_1,\tilde{k}_2);(k_1,k_2)\right)$$

$$\Leftrightarrow \quad \underbrace{\theta(b_1-(b_2-(\ell_2-1)),\ell_2-1)}_{=P(Y_1=b_1-(b_2-(\ell_2-1)),Y_2=\ell_2-1)}\cdot\frac{1}{2}v \quad \underbrace{\sum_{k_1=b_1-(b_2-\ell_2)}^{b_1}\theta(k_1,\ell_2-1)}_{=P(Y_1>b_1-(b_2-(\ell_2-1)),Y_2=\ell_2-1)}\cdot v$$

$$=\underbrace{\sum_{\tilde{k}_1=0}^{b_1}\theta(\tilde{k}_1,\ell_2)}_{=P(Y_2=\ell_2)}\cdot\lambda_2.$$

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

$$P(Y_2=\ell_2)\cdot\lambda_2 = P(Y_1=b_1-(b_2-(\ell_2-1)),Y_2=\ell_2-1)\cdot\frac{1}{2}v$$
$$+ P(Y_1>b_1-(b_2-(\ell_2-1)),Y_2=\ell_2-1)\cdot v$$
$$= P(Y_1=b_1-b_2+\ell_2-1,Y_2=\ell_2-1)\cdot\frac{1}{2}v$$
$$+ P(Y_1>b_1-b_2+\ell_2-1,Y_2=\ell_2-1)\cdot v.$$

$\square$

6 Conclusion

The paper investigated a supply chain where routing of items depends on the on-hand inventory at the locations with the aim to obtain "load balancing". The main result of the paper is to prove that the stationary distribution has a product form. This means 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. For the marginal distribution of the production subsystem an explicit solution was derived and for the marginal distribution of the inventory-replenishment subsystem an explicit solution or a recursive algorithm were deduced for some special cases.

To extend the results to queueing-inventory systems with other replenishment policies is part of our ongoing research. First results for queueing-inventory systems with (r, S)-policy with location specific items can be found in Otten (2017, Chapter 7).

Data availability The author does not analyse or generate any datasets, because the work proceeds within a theoretical and mathematical approach.

Declarations

Conflict of interest The author declares that there is no conflict of interest.

Appendix A Computational example

Motivated by the computational example in Chow and Kohler (1979, Appendix, pp. 360f.) we consider an example with two locations, base stock levels $b_1 = 4$, $b_2 = 3$ and arrival rates $\lambda_1, \lambda_2 > 0$. The state transition diagram is presented in Fig. 6.

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 solution steps for this system using the algorithm follow:

▶ Set $\widetilde{\theta}(4, 0) = 1$.
▶ For $k_2 = 3$:

(1)(a) Set $\widetilde{\theta}(0, 3) = \kappa$.
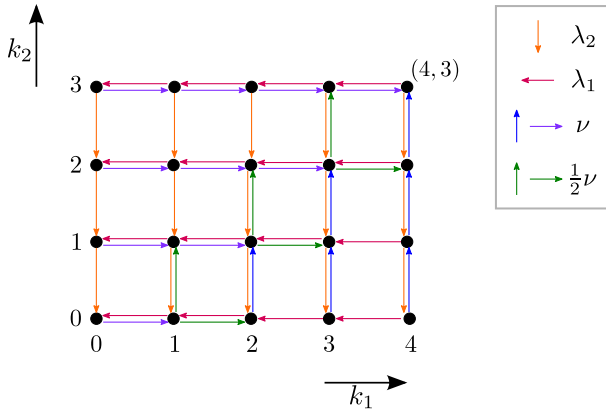
Fig. 6 State transition diagram of a system with two locations and $b_1 = 4$, $b_2 = 3$ and $\lambda_1, \lambda_2 > 0$

(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

$$\widetilde{\theta}(4, 1) = \frac{\lambda_1 + \nu}{\lambda_2} =: c^{(4,1)}(\lambda_1, \lambda_2, \nu).$$

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

$$\widetilde{\theta}(4, 2) = \frac{(\lambda_1 + \lambda_2 + \nu)}{\lambda_2} \cdot c^{(4,1)}(\lambda_1, \lambda_2, \nu) =: c^{(4,2)}(\lambda_1, \lambda_2, \nu).$$

(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

$$\widetilde{\theta}(2, 3) = \frac{\widetilde{\theta}(1, 3) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(0, 3) \cdot \nu}{\lambda_1}$$

$$= \frac{c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - \nu}{\lambda_1} \cdot \kappa =: c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

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

$$\widetilde{\theta}(3, 3) = \frac{\widetilde{\theta}(2, 3) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1, 3) \cdot \nu}{\lambda_1}$$

$$= \frac{c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(1,3)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1} \cdot \kappa$$

$$=: c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

(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

$$\widetilde{\theta}(4, 3) = \frac{\widetilde{\theta}(3, 3) \cdot \nu + \widetilde{\theta}(4, 2) \cdot \nu}{\lambda_1 + \lambda_2} = \frac{c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1 + \lambda_2} \cdot \kappa + \frac{c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\lambda_1 + \lambda_2}$$

$$=: c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(4,3)}(\lambda_1, \lambda_2, \nu).$$

(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

$$\widetilde{\theta}(3, 2)$$

$$= \frac{\widetilde{\theta}(3, 3) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(4, 3) \cdot \lambda_1 - \widetilde{\theta}(2, 3) \cdot \nu}{\frac{1}{2} \cdot \nu}$$

$$= \frac{c^{(3,3)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_1 - c^{(2,3)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\frac{1}{2} \cdot \nu} \cdot \kappa$$

$$- \frac{c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_1}{\frac{1}{2} \cdot \nu}$$

$$=: c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(3,2)}(\lambda_1, \lambda_2, \nu).$$

(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

$$c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu)$$

$$= \left(c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \right) \cdot \lambda_2$$

$$+ \left(c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(3,2)}(\lambda_1, \lambda_2, \nu) \right) \cdot \frac{1}{2} \cdot \nu + c^{(4,1)}(\lambda_1, \lambda_2, \nu) \cdot \nu$$

$$\Leftrightarrow c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \lambda_2 + c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa \cdot \frac{1}{2} \cdot \nu$$

$$= c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(4,1)}(\lambda_1, \lambda_2, \nu) \cdot \nu$$

$$- c_2^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2} \cdot \nu - c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_2$$

$$\Leftrightarrow \kappa = \left[c^{(4,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c^{(4,1)}(\lambda_1, \lambda_2, \nu) \cdot \nu \right.$$

$$\left. - c_2^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2} \cdot \nu - c_2^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_2 \right]$$

$$\cdot \frac{1}{c_1^{(4,3)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_2 + c_1^{(3,2)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2} \cdot \nu}$$

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

$$\widetilde{\theta}(1, 2) = \frac{\widetilde{\theta}(0, 3) \cdot \lambda_2 - \widetilde{\theta}(0, 2) \cdot (\lambda_2 + \nu)}{\lambda_1}$$

$$= \frac{\widetilde{\theta}(0, 3) \cdot \lambda_2}{\lambda_1} - \frac{(\lambda_2 + \nu)}{\lambda_1} \cdot \kappa =: c_1^{(1,2)}(\lambda_1, \lambda_2) + c_2^{(1,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

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

$$\widetilde{\theta}(2, 2) = \frac{\widetilde{\theta}(1, 2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1, 3) \cdot \lambda_2 - \widetilde{\theta}(0, 2) \cdot \nu}{\lambda_1}$$

$$= \frac{c_1^{(1,2)}(\lambda_1, \lambda_2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(1, 3) \cdot \lambda_2}{\lambda_1}$$

$$+ \frac{\left[c_2^{(1,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - \nu \right]}{\lambda_1} \cdot \kappa$$

$$=: c_1^{(2,2)}(\lambda_1, \lambda_2, \nu) + c_2^{(2,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

(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

$$\widetilde{\theta}(2, 1)$$
$$= \frac{\widetilde{\theta}(2, 2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3, 2) \cdot \lambda_1 - \widetilde{\theta}(2, 3) \cdot \lambda_2 - \widetilde{\theta}(1, 2) \cdot \nu}{\frac{1}{2}\nu}$$
$$= \frac{c_1^{(2,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3, 2) \cdot \lambda_1 - \widetilde{\theta}(2, 3) \cdot \lambda_2 - c_1^{(1,2)}(\lambda_1, \lambda_2) \cdot \nu}{\frac{1}{2}\nu}$$
$$+ \frac{c_2^{(2,2)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - c_2^{(1,2)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\frac{1}{2}\nu} \cdot \kappa$$
$$=: c_1^{(2,1)}(\lambda_1, \lambda_2, \nu) + c_2^{(2,1)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

(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

$$\widetilde{\theta}(3, 1) = \frac{\widetilde{\theta}(3, 2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(4, 2) \cdot \lambda_1 - \widetilde{\theta}(3, 3) \cdot \lambda_2 - \widetilde{\theta}(2, 2) \cdot \nu}{\nu}$$
$$= \frac{\widetilde{\theta}(3, 2) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(4, 2) \cdot \lambda_1 - \widetilde{\theta}(3, 3) \cdot \lambda_2 - c_1^{(2,2)}(\lambda_1, \lambda_2, \nu) \cdot \nu}{\nu}$$
$$- c_2^{(2,2)}(\lambda_1, \lambda_2, \nu) \cdot \kappa$$
$$=: c_1^{(3,1)}(\lambda_1, \lambda_2, \nu) + c_2^{(3,1)}(\lambda_1, \lambda_2, \nu) \cdot \kappa.$$

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

$$\widetilde{\theta}(3, 0)$$
$$= \frac{\widetilde{\theta}(3, 1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(4, 1) \cdot \lambda_1 - \widetilde{\theta}(3, 2) \cdot \lambda_2 - \widetilde{\theta}(2, 1) \cdot \frac{1}{2}\nu}{\nu}$$
$$= \frac{c_1^{(3,1)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(4, 1) \cdot \lambda_1 - \widetilde{\theta}(3, 2) \cdot \lambda_2 - c_1^{(2,1)}(\lambda_1, \lambda_2, \nu)\frac{1}{2}\nu}{\nu}$$
$$+ \frac{c_2^{(3,1)}(\lambda_1, \lambda_2, \nu) \cdot \kappa(\lambda_1 + \lambda_2 + \nu) - c_2^{(2,1)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2}\nu}{\nu} \cdot \kappa$$

$$=: c_1^{(3,0)}(\lambda_1, \lambda_2, v) + c_2^{(3,0)}(\lambda_1, \lambda_2, v) \cdot \kappa.$$

(d) Use the GBE of state $(3, 0)$ to solve for κ
($\widetilde{\theta}(4, 0)$ is known independent of κ from step (1)):

$$\widetilde{\theta}(3, 0) \cdot (\lambda_1 + v) = \widetilde{\theta}(4, 0) \cdot \lambda_1 + \widetilde{\theta}(3, 1) \cdot \lambda_2,$$

which is equivalent to

$$\left(c_1^{(3,0)}(\lambda_1, \lambda_2, v) + c_2^{(3,0)}(\lambda_1, \lambda_2, v) \cdot \kappa \right) \cdot (\lambda_1 + v)$$

$$= \widetilde{\theta}(4, 0) \cdot \lambda_1 + \left(c_1^{(3,1)}(\lambda_1, \lambda_2, v) + c_2^{(3,1)}(\lambda_1, \lambda_2, v) \cdot \kappa \right) \cdot \lambda_2$$

$$\Leftrightarrow c_2^{(3,0)}(\lambda_1, \lambda_2, v) \cdot \kappa \cdot (\lambda_1 + v) - c_2^{(3,1)}(\lambda_1, \lambda_2, v) \cdot \kappa \cdot \lambda_2$$

$$= \widetilde{\theta}(4, 0) \cdot \lambda_1 + c_1^{(3,1)}(\lambda_1, \lambda_2, v) \cdot \lambda_2 - c_1^{(3,0)}(\lambda_1, \lambda_2, v) \cdot (\lambda_1 + v)$$

$$\Leftrightarrow \kappa = \left[\widetilde{\theta}(4, 0) \cdot \lambda_1 + c_1^{(3,1)}(\lambda_1, \lambda_2, v) \cdot \lambda_2 - c_1^{(3,0)}(\lambda_1, \lambda_2, v) \cdot (\lambda_1 + v) \right]$$

$$\cdot \frac{1}{c_2^{(3,0)}(\lambda_1, \lambda_2, v) \cdot (\lambda_1 + v) - c_2^{(3,1)}(\lambda_1, \lambda_2, v) \cdot \lambda_2}$$

if the denominator is not equal to zero.

(e) Substitute the value of κ into the equations in the above steps (2)(a)–(c).

(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)):

$$\widetilde{\theta}(0, 1) \cdot (\lambda_2 + v) = \widetilde{\theta}(0, 2) \cdot \lambda_2 + \widetilde{\theta}(1, 1) \cdot \lambda_1,$$

which is equivalent to

$$\widetilde{\theta}(1, 1) = \frac{(\lambda_2 + v)}{\lambda_1} \cdot \kappa - \frac{\widetilde{\theta}(0, 2) \cdot \lambda_2}{\lambda_1} =: c_1^{(1,1)}(\lambda_1, \lambda_2) \cdot \kappa + c_2^{(1,1)}(\lambda_1, \lambda_2).$$

(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)):

$$\widetilde{\theta}(1, 1) \cdot (\lambda_1 + \lambda_2 + v) = \widetilde{\theta}(2, 1) \cdot \lambda_1 + \widetilde{\theta}(1, 2) \cdot \lambda_2 + \widetilde{\theta}(0, 1) \cdot v + \widetilde{\theta}(1, 0) \cdot \frac{1}{2} v,$$

which is equivalent to

$$\widetilde{\theta}(1, 0) = \frac{\widetilde{\theta}(1, 1) \cdot (\lambda_1 + \lambda_2 + v) - \widetilde{\theta}(2, 1) \cdot \lambda_1 - \widetilde{\theta}(1, 2) \cdot \lambda_2 - \widetilde{\theta}(0, 1) \cdot v}{\frac{1}{2} v}$$

$$= \frac{c_1^{(1,1)}(\lambda_1, \lambda_2) \cdot (\lambda_1 + \lambda_2 + v) - v}{\frac{1}{2} v} \cdot \kappa$$

$$+ \frac{c_2^{(1,1)}(\lambda_1, \lambda_2) \cdot (\lambda_1 + \lambda_2 + v) - \widetilde{\theta}(2, 1) \cdot \lambda_1 - \widetilde{\theta}(1, 2) \cdot \lambda_2}{\frac{1}{2} v}$$

$$=: c_1^{(1,0)}(\lambda_1, \lambda_2, v) \cdot \kappa + c_2^{(1,0)}(\lambda_1, \lambda_2, v).$$

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)):

$$\widetilde{\theta}(2, 1) \cdot (\lambda_1 + \lambda_2 + v) = \widetilde{\theta}(3, 2) \cdot \lambda_1 + \widetilde{\theta}(2, 2) \cdot \lambda_2 + \widetilde{\theta}(2, 0) \cdot v + \widetilde{\theta}(1, 1) \cdot v,$$

which is equivalent to

$$
\begin{aligned}
\widetilde{\theta}(2,0) &= \frac{\widetilde{\theta}(2,1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3,2) \cdot \lambda_1 - \widetilde{\theta}(2,2) \cdot \lambda_2 - \widetilde{\theta}(1,1) \cdot \nu}{\nu} \\
&= -\frac{c_1^{(1,1)}(\lambda_1, \lambda_2) \cdot \nu}{\nu} \cdot \kappa \\
&\quad + \frac{\widetilde{\theta}(2,1) \cdot (\lambda_1 + \lambda_2 + \nu) - \widetilde{\theta}(3,2) \cdot \lambda_1 - \widetilde{\theta}(2,2) \cdot \lambda_2 - c_2^{(1,1)}(\lambda_1, \lambda_2) \cdot \nu}{\nu} \\
&=: c_1^{(2,0)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(2,0)}(\lambda_1, \lambda_2, \nu).
\end{aligned}
$$

(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

$$
\begin{aligned}
\widetilde{\theta}(0,0) &= \frac{\widetilde{\theta}(1,0) \cdot (\lambda_1 + \nu) - \widetilde{\theta}(2,0) \cdot \lambda_1 - \widetilde{\theta}(1,1) \cdot \lambda_2}{\nu} \\
&= \frac{c_1^{(1,0)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \nu) - c_1^{(2,0)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_1 - c_1^{(1,1)}(\lambda_1, \lambda_2) \cdot \lambda_2}{\nu} \cdot \kappa \\
&\quad + \frac{c_2^{(1,0)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \nu) - c_2^{(2,0)}(\lambda_1, \lambda_2, \nu) \cdot \lambda_1 - c_2^{(1,1)}(\lambda_1, \lambda_2) \cdot \lambda_2}{\nu} \\
&=: c_1^{(0,0)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(0,0)}(\lambda_1, \lambda_2, \nu).
\end{aligned}
$$

(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)):

$$
\begin{aligned}
&\widetilde{\theta}(2,0) \cdot (\lambda_1 + \nu) = \widetilde{\theta}(3,0) \cdot \lambda_1 + \widetilde{\theta}(2,1) \cdot \lambda_2 + \widetilde{\theta}(1,0) \cdot \frac{1}{2}\nu \\
&\Leftrightarrow \left(c_1^{(2,0)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(2,0)}(\lambda_1, \lambda_2, \nu) \right) \cdot (\lambda_1 + \nu) \\
&= \widetilde{\theta}(3,0) \cdot \lambda_1 + \widetilde{\theta}(2,1) \cdot \lambda_2 \quad + \left(c_1^{(1,0)}(\lambda_1, \lambda_2, \nu) \cdot \kappa + c_2^{(1,0)}(\lambda_1, \lambda_2, \nu) \right) \cdot \frac{1}{2}\nu \\
&\Leftrightarrow \kappa \cdot \left(c_1^{(2,0)}(\lambda_1, \lambda_2, \nu) - c_1^{(1,0)}(\lambda_1, \lambda_2, \nu) \right) \\
&= \widetilde{\theta}(3,0) \cdot \lambda_1 + \widetilde{\theta}(2,1) \cdot \lambda_2 - c_2^{(2,0)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \nu) + c_2^{(1,0)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2}\nu \\
&\Leftrightarrow \kappa \\
&= \frac{\widetilde{\theta}(3,0) \cdot \lambda_1 + \widetilde{\theta}(2,1) \cdot \lambda_2 - c_2^{(2,0)}(\lambda_1, \lambda_2, \nu) \cdot (\lambda_1 + \nu) + c_2^{(1,0)}(\lambda_1, \lambda_2, \nu) \cdot \frac{1}{2}\nu}{c_1^{(2,0)}(\lambda_1, \lambda_2, \nu) - c_1^{(1,0)}(\lambda_1, \lambda_2, \nu)}
\end{aligned}
$$

if the denominator is not equal to zero.

(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

$$
\widetilde{\theta}(k_1, k_2) \leftarrow \frac{\widetilde{\theta}(k_1, k_2)}{\sum_{k_1=0}^{b_1} \sum_{k_2=0}^{b_2} \widetilde{\theta}(k_1, k_2)}
$$

References

Abouee-Mehrizi, H., Baron, O., & Berman, O. (2014). Exact analysis of capacitated two-echelon inventory systems with priorities. *Manufacturing & Service Operations Management, 16*(4), 561–577. https://doi.org/10.1287/msom.2014.0494.

Axsäter, S. (2000). *Inventory control*. Kluwer Academic Publishers.

Bijvank, M., & Vis, F. A. (2011). Lost-sales inventory theory: A review. *European Journal of Operations Research, 215,* 1–13. https://doi.org/10.1016/j.ejor.2011.02.004.

Chao, X., Miyazawa, M., & Pinedo, M. (1999). *Queueing networks—Customers, signals, and product form solutions*. Wiley.

Chow, Y. C., & Kohler, W. H. (1979). Models for dynamic load balancing in a heterogeneous multiple processor system. *IEEE Transactions on Computers, 28*(5), 354–361. https://doi.org/10.1109/TC.1979.1675365.

Chow, Y. C., & Kohler, W. H. (1977). Dynamic load balancing in homogeneous two-processor distributed systems. In K. M. Chandy & M. Reiser (Eds.), *Computer Performance* (pp. 39–52). North-Holland.

Daduna, H. (1985). The cycle-time distribution in a central server network with state-dependent branching. *Optimization, 16,* 617–626. https://doi.org/10.1080/02331938508843057.

Daduna, H. (1987). Cycle times in a starlike network with state-dependent routing. *Journal of Applied Mathematics and Simulation, 1,* 1–12. https://doi.org/10.1155/S1048953388000012.

Daduna, H. (1990). Exchangeable items in repair systems: Delay times. *Operations Research, 38*(2), 349–354. https://doi.org/10.1287/opre.38.2.349.

Donselaar, K. H. V., & Broekmeulen, R. A. C. M. (2013). 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. https://doi.org/10.1016/j.ijpe.2011.05.020.

Flatto, L., & McKean, H. P. (1977). Two queues in parallel. *Communications on Pure and Applied Mathematics, 30*(2), 255–263. https://doi.org/10.1002/cpa.3160300206.

Gruen, T. W., Corsten, D., & Bharadwaj, S. (2002). *Retail out-of-stocks: A worldwide examination of extent causes and consumer responses*. Grocery Manufacturers of America.

Herzog, U., Woo, L., & Chandy, K. M. (1975). Solution of queueing problems by a recursive technique. *IBM Journal of Research and Development, 19*(3), 295–300. https://doi.org/10.1147/rd.193.0295.

Hordijk, A., & Dijk, N. M. V. (1984). Networks of queues. Part I, II. Lecture notes in control and information sciences. In F. Baccelli & G. Fayolle (Eds.), *Modelling and performance evaluation methodology* (Vol. 60, pp. 151–205). Springer. https://doi.org/10.1007/BFb0005165.

Jackson, J. R. (1957). Networks of waiting lines. *Operations Research, 5,* 518–521. https://doi.org/10.1287/opre.5.4.518.

Kelly, F. P. (1979). *Reversibility and stochastic networks*. Wiley.

Krenzler, R., Daduna, H., & Otten, S. (2016). Jackson networks in non-autonomous random environments. *Advances in Applied Probability*. https://doi.org/10.1017/apr.2016.2.

Krishnamoorthy, A., Shajin, D., & Narayanan, V. C. (2021). Inventory with positive service time: A survey. In V. Anisimov & N. Limnios (Eds.), *Queueing theory* (2nd ed., Vol. 6, pp. 201–237). Wiley. https://doi.org/10.1002/9781119755234.ch6.

Morse, P. M. (1958). *Queues, inventories and maintenance*. Wiley. https://doi.org/10.2307/2342909.

Otten, S. (2017). Integrated models for performance analysis and optimization of queueing-inventory-systems in logistic networkst. PhD thesis, University of Hamburg, Department of Mathematics.

Otten, S., Krenzler, R., & Daduna, H. (2016). Models for integrated production-inventory systems: Steady state and cost analysis. *International Journal of Production Research, 54*(20), 6174–6191. https://doi.org/10.1080/00207543.2015.1082669.

Pittel, B. (1979). Closed exponential networks of queues with saturation: The Jackson-type stationary distribution and its asymptotic analysis. *Mathematics of Operations Research, 4,* 357–378. https://doi.org/10.1287/moor.4.4.357.

Ravid, R., Boxma, O. J., & Perry, D. (2013). Repair systems with exchangeable items and the longest queue mechanism. *Queueing Systems*. https://doi.org/10.1007/s11134-012-9319-5.

Reed, J., & Zhang, B. (2017). Managing capacity and inventory for multi-server make-to-stock queues. *Queueing Systems, 86,* 61–94. https://doi.org/10.1007/s11134-017-9519-0.

Rego, J. R. D., & Mesquita, M. A. D. (2011). Spare parts inventory control: A literature review. *Produção, 21*(4), 656–666. https://doi.org/10.1590/S0103-65132011005000002.

Rubio, R., & Wein, L. M. (1996). Setting base stock levels using product-form queueing networks. *Management Science, 42,* 259–268. https://doi.org/10.1287/mnsc.42.2.259.

Schassberger, R. (1984). Decomposable stochastic networks: Some observations. Lecture notes in control and information sciences. In F. Baccelli & G. Fayolle (Eds.), *Modelling and performance evaluation methodology* (Vol. 60, pp. 137–150). Springer. https://doi.org/10.1007/BFb0005173.

Silver, E. A., Pyke, D., & Peterson, R. (1998). *Inventory management and production planning and scheduling* (3rd ed.). Wiley.

Tempelmeier, H. (2005). Bestandsmanagement in Supply Chains. Books on Demand.

Towsley, D. (1980). Queueing network models with state-dependent routing. *Journal of the Association for Computing Machinery, 27*(2), 323–337. https://doi.org/10.1145/322186.322196.

Verhoef, P., & Sloot, L. M. (2006). Out-of-stock: Reactions, antecedents, management solutions, and a future perspective. In M. Krafft & M. K. Mantrala (Eds.), *Retailing in the 21st century: Current and future trends* (pp. 239–253). Springer. https://doi.org/10.1007/978-3-540-72003-4_18.

Zazanis, M. (1994). Push and pull systems with external demands. In: *Proceedings of the 32nd Allerton conference on communication, control, and computing, Allerton.*

Zipkin, P. H. (2000). *Foundations of inventory management*. McGraw-Hill.