Queueing Systems in a Random Environment

Dissertation zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Mathematik
der Universität Hamburg
vorgelegt von Ruslan Krenzler
Hamburg, 2016

Tag der Disputation: 13. Juli 2016
Folgende Gutachter empfehlen die Annahme der Dissertation:
Prof. Dr. Hans Daduna
Prof. Dr. Bernd Heidergott
Prof. Dr. Udo R. Krieger

Für Serik und Erika

Contents

Acknowledgments

I would like to express my gratitude to my advisor Hans Daduna for his extraordinary support and encouragement. His expert advice in mathematics, didactics, and sciences in general was very valuable. It was and is a great pleasure to work with him on joint articles, several of them became an essential part of this thesis.

I would like to thank Bernd Heidergott and Udo R. Krieger for reviewing this thesis.

I would like to thank Hendrik Baumann for the interesting discussions about matrix-geometric solutions and stochastic processes in continuous state space. I would like to thank Sabine Le Borne for her hints about essentially dominant matrices and their invertibility properties. I would like to thank Achyutha Krishnamoorthy and Manikandan Rangaswamy for the fruitful discussions about product-form results in inventory models. I would like to thank Evsey Morozov and Alexander Rumyantsev for the collaboration in developing product-form models of desktop grid and volunteer computing networks. I would like to thank Dietrich Stoyan for introducing me to the application field of mining industry.

I would like to thank my colleague and coauthor Sonja Otten for the excellent cooperation and friendly working atmosphere.

I would like to thank the anonymous reviewers for their comments on the parts of this thesis which were submitted to journals and conferences.

I would like to thank Jens Struckmeier whose lectures about mathematical modeling for pupils and students had strong impact on my research focus and research methods.

I would like to thank my family for their great support during the time when I was writing my thesis.

Introduction

Queueing networks with product-form steady-state distribution have found many fields of applications, e.g. production systems, telecommunications, and computer system modeling. The success of this class of models and its relatives stems from the simple structure of the steady-state distribution which provides access to easy performance evaluation procedures. Starting from the work of Jackson [Jac57] various generalizations have been developed.

In real world queueing systems are not isolated and interact with their environment. Adding a random environment to a model usually makes the model more realistic but also more complex to analyze. Nevertheless, under some conditions it is still possible to obtain analytical results. A branch of research which recently has found interest are queueing networks in a random environment with product form steady-state distributions.

In this thesis we develop a general theory that comprise models with stationary product-form distribution in inventory theory in [Sch04] and Jackson networks with unreliable nodes with stationary product-form distribution in [Sau06]. An important property of the resulting general model is that the queueing system and the environment interact in both directions: the queues can influence the environment and the environment influences the queues.

In Part I we analyze single-queue systems. In Section 1 we introduce a *loss system*. In Section 2 we generalize product form lost-sales inventory models from [Sch04] and several other published papers with related models as a loss system with exponential service time. The term *loss* means that customers get lost when the environment stays in some special states – the blocking states. In Section 2.1.4 we develop an approximation method for system without loss of customers based on loss systems. In Section 2.2 we apply our loss system results in fields different from inventory management: we analyze in detail an unreliable server with preventive maintenance in Section 2.2.4, a node of a wireless sensor network in Section 2.2.5, and a crusher station in open-pit mining in Section 2.2.6.

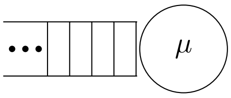
In Section 3 we analyze the Markov chain embedded at departure instants of the loss system. The embedded Markov chains are an important tool for analyzing queueing system with general service times – the $M/G/1/\infty$ queues. The famous and frequently used result in classical $M/G/1/\infty$ theory is that the steady-state distribution of an $M/G/1/\infty$ system as continuous time process and as embedded Markov chain, observed at departure times, are the same. We show that this is in general not true for the steady-state distribution of loss systems. We use an embedded Markov chain analysis to extend our results from Section 2 to some loss systems with general service times.

In Part II we extend our results for a single-queue loss system to Jackson networks in a random environment. We replace the concept of loss of customers by special rerouting regimes. We establish a connection between these rerouting regimes and randomized random walks. In Section 8 we consider systems where the interaction between environment and queuing system depend on the number of customers in the system. This extension finally allows us to include results about Jackson networks with unreliable nodes from [Sau06] as special cases.

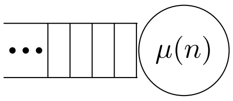
Notations and conventions:

- $\mathbb{R}_0^+ = [0, \infty)$, $\mathbb{R}^+ = (0, \infty)$, $\mathbb{N}$=$\{1,2,3,\ldots\}$, $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$.
- $\mathbb{K}$ is a field, either $\mathbb{R}$ or $\mathbb{C}$.
- A value is said to be *positive* if it is greater than zero. A value said to be *negative* if it less than zero. Sometimes we will use the terms *strictly positive* or *strictly negative* to emphasizes that the corresponding values are not zero. We call a value *non-negative* if it is greater than or equal to zero.
- $A \subset B$ means A is a subset of B. $A \subsetneq B$ means A is a proper subset of B.
- We write $C = A \uplus B$ to emphasize that C is the union of disjoint sets A and B.
- $1_{[expression]}$ is the indicator function which is 1 if *expression* is true and 0 otherwise.
- $\delta_{xy} := 1_{[x=y]}$ is the Kronecker-Delta.
- For any quadratic matrix V we define $\text{diag}(V)$ as the matrix with the same diagonal as V, while all other entries are 0.
- $\mathbf{e}$ is a vector whose all elements are 1. The dimension of $\mathbf{e}$ will always be clear from the context.
- Empty sums are 0 and empty products are 1.
- 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 equal zero. From this definition 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 or equal one.
- All random variables and processes occurring henceforth are defined on a common underlying probability space $(\Omega, \mathcal{F}, P)$.
- We call a process in continuous time *regular* if it is non-explosive, its transition intensity matrix is conservative (row sums are 0) and all diagonal elements of the transition intensity matrix are finite.
- In this thesis, we assume that all stochastic processes in continuous time are regular and that a version with right continuous paths with left limits (cadlag) is given.
- If we use $\cdot/\cdot/1/N$ notation for queueing systems with one server and N waiting places, the maximal number of customers in this queuing system is $N+1$.
- In diagrams, we use three different symbolic representations of queues shown in Figure 0.2 on page x.

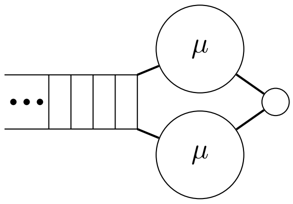
- In diagrams, we use big circles to represent sever of the queues, a queue can have multiple servers. In the circle we write the service rate of each server. See Figure 0.1 on page x.



(a) $\cdot/\cdot/1/\infty$ queue with constant service rate μ

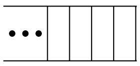


(b) $\cdot/\cdot/1/\infty$ queue with service rate $\mu(n)$ depending on number of customers in the system

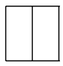


(c) $\cdot/\cdot/2/\infty$ queue with two servers where each server has constant service rate μ

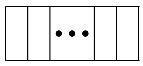
Figure 0.1.: Different symbolic representations of servers at a queue with infinite number of waiting places



(a) queue with infinite number of waiting places



(b) queue with two waiting places



(c) queue with finite but unknown or large number of places

Figure 0.2.: Different symbolic representation of the queues.

Part I.

Birth-death-loss processes in a random environment

1. Introduction

Product form networks of queues are common models for easy to perform structural and quantitative first order analysis of complex networks in Operations Research applications. The most prominent representatives of this class of models are the Jackson [Jac57] and Gordon-Newell [GN67] networks and their generalizations as BCMP [BCMP75] and Kelly [Kel76] networks, for a short review see [Dad01].

Standard mathematical description of this class of models is by time homogeneous Markovian vector processes, where each coordinate represents the behaviour of one of the queues. Product form networks are characterized by the fact that in steady state (at any fixed time t) the joint distribution of the multi-dimensional (over nodes) queueing process is the product of the stationary marginal distributions of the individual nodes' (non Markovian) queueing processes. With respect to the research described in this note the key point is that the coordinates of the vector process represent objects of the same class, namely queueing systems.

In Operations Research applications queueing systems constitute an important class of models in very different settings. Nevertheless, in many applications those parts of, e.g., a complex production system which are modeled by queues interact with other subsystems which usually can not be modeled by queues. We will describe two prototype situations which will be considered in detail throughout the paper as introductory examples. These examples deal with interaction of (i) a queueing system with an inventory, and (ii) a queueing system with its environment, which influences the availability of the server.

Introductory example (i): Production-inventory system. Typically, there is a manufacturing system, (machine, modeled by a queueing system) which assembles delivered raw material to a final product, consuming in the production process some further material (we will call these additional pieces of material "items" henceforth) which is hold in inventories.

Introductory example (ii): Availability of a production system. The manufacturing system (machine, modeled by a queueing system) may break down caused by influences of its environment or by wear out of its server and has to be repaired.

Our present research is motivated by the observation that in both situations we have to construct an integrated model with a common structure:

- A production processes modeled by queueing systems and
- an additional relevant part of the system with different character, e.g., inventory control or availability control.

In both models the components strongly interact, and although the additional feature are quite different, we will extract similarities. This motivates to construct a unified model which encompasses both introductory examples and as we will show, many other examples in different fields. For taking a general nomenclature we will subsume in any case the "additional relevant part of the system" attached to the queueing model as "environment of the queue".

Our construction of "queueing systems in a random environment" will result in a set of product form stationary distributions for the Markovian joint queueing-environment process, i.e., in equilibrium the coordinates at fixed time points decouple: The stationary distribution of the joint queueing-environment process is the product of stationary marginal distributions of the queue and the environment, which in general cannot be described as a Markov process of their own. The key point is that the coordinates of the vector process represent very different classes of objects.

Product form stationary distributions for the introductory examples have been found only recently.

(i) Schwarz, Sauer, Daduna, Kulik, and Szekli [SSD+06] discovered product forms for the steady-state distributions of an $M/M/1/\infty$ under standard order policies with lost sales. Further contributions to product form results in this field are by Vineetha [Vin08], Saffari, Haji, and Hassanzadeh [SHH11], and Saffari, Asmussen, and Haji [SAH13]. An early paper of Berman and Kim [BK99] can be considered to contribute to integrated models with product form steady state.

(ii) For classical product form networks of queues in [SD03a] the influence of breakdown and repair of the nodes was studied and it was proved that under certain conditions a product form equilibrium for such networks of unreliable servers exists. The Markovian description of the system encompasses coordinates describing (possibly many) queues and an additional coordinate to indicate the reliability status of the system, for more details see [Sau06].

Related research on queueing networks in a random environment is by Zhu [Zhu94], Economou [Eco03, Eco04, Eco05], Tsitsiashvili, Osipova, Koliev, Baum [TOKB02], and Balsamo, Marin [BM13]. There usually the environment is a Markov process of its own, which is the case neither in our model nor in the motivating results in [SSD+06] and [SD03a], which consider our introductory examples. A queueing network in a random environment can be modelled as a quasi birth-death (QBD) process whose level describes the length of one particular queue and whose phase describes the environment state and the state of the other queues. In [LR99, Chapter 15], Latouche and Ramaswami and, in [RT96], Ramaswami and Taylor analyzed product form QBD. They derived necessary and sufficient conditions for the product-form steady-state distribution and applied the results to various product-form queueing systems.

An important common aspect of the interaction in both introductory examples leads to the term **loss system** for our general interacting system: Whenever for the queue, respectively,

- the inventory is depleted (i),
- the machine is broken down (ii),

service at the production unit is interrupted due to stock out (i), resp. no production capacity available (ii). Additionally, during the time the interruption continues no new arrivals are admitted to both systems, due to lost sales and because customers prefer to enter some other working server.

Note, that this loss of customers is different from what is usually termed loss systems in pure queueing theory, where loss of customers happens, when the finite waiting space is filled up to maximal capacity.

Following the above description, in our present investigation of complex systems we always start with a queueing system as one subsystem and a general attached other subsystem (the environment) which imposes side constraints on the queueing process and in general interacts in both directions with the queue. In typical cases there will be a part of the environment's state space, the states of which we shall call "blocking states", with the following property: Whenever the environment enters a blocking state, the service process will be interrupted and no new arrivals are admitted to enter the system and are lost to the system forever.

The interaction of the components in this class of models is that jumps of the queue may enforce the environment to jump instantaneously, and in the other direction the evolving environment may interrupt service and arrivals at the queue, by entering blocking states, and when leaving the set of blocking states service is resumed and new customers are admitted again.

We describe our exponential system in Section 2.1.1 and start our detailed investigation in Section 2.1.1.1. Our main result (Theorem 2.1.5) is that, although production and environment strongly interact, asymptotically and in equilibrium (at fixed time instants) the production process and the environment process seem to decouple, which means that a product form equilibrium emerges.

This shows that the mentioned independence results in [SSD$^+$06], [SD03a], [SHH11], [SAH13], and [Vin08] do not depend on the specific properties of the attached second subsystem. Furthermore, we will show that the theorem can be interpreted as a strong insensitivity property of the system: As long as ergodicity is maintained, the environment can change drastically without changing the steady-state distribution of the queue length.

And, vice versa, it can be seen that the environment's steady state will not change when the service capacity of the production will change.

We shall discuss this with related problems and some complements to the theorem in more detail in Section 2.1.1.1 after presenting our main result there. In Section 2.1.2 we investigate the case of finite waiting space at the queue, so in the classical loss system we introduce additional losses due to the impact of the environment on the production process. Astonishingly, there occur new structural problems when product form steady states are found.

In Section 2.2 we present a bulk of applications of our abstract modeling process to systems found in the literature. We show especially, that our main theorem allows to generalize rather directly many of the previous results. In Section 3.1 we consider the systems (which live in continuous time) at departure instants only, which results in considering an embedded Markov chain. We find that the behaviour of the embedded Markov chain is often considerably different from that of the original continuous time Markov process investigated in Section 2. Especially, it is a non trivial task to decide whether the stationary distribution of the embedded Markov chain (at departure instants) is of product form as well.

For exponential queueing systems we show that there is a product form equilibrium under rather general conditions. We provide this stationary distributions explicitly in

Theorem 3.1.16, showing that the marginal queue length distribution of the embedded chain is the same as in continuous time, and discuss the relation relation between the respective marginal environment distributions, which are not equal but related by a transformation which we determine explicitly.

To emphasize the problems arising from the interaction of the two components of integrated systems, we remind the reader, that for ergodic $M/M/1/\infty$ queues the limiting and stationary distribution of the continuous time queue length process and the Markov chains embedded at either arrival instants or departure instants are the same. In connection with this, we revisit some of Vineetha's [Vin08] queueing-inventory systems, using similarly embedded Markov chain techniques.

A striking observation is moreover, that for a system which is ergodic in the continuous time Markovian description the Markov chain embedded at departure instants may be not ergodic. The reason for this is two-fold. Firstly, the embedded Markov chain may have inessential states due to the specified interaction rules. Secondly, even when we delete all inessential states, the resulting single positive recurrent class may be periodic. We study this problem in depth in Section 3.1.2.

In Section 3.2 we show that for systems with non-exponential service times more restrictive constraints are needed, which we prove by a counter example where the environment represents an inventory attached to an $M/D/1$ queue. Such integrated queueing-inventory systems are dealt with in the literature previously, e.g. in [Vin08]. Further applications are, e.g., in modeling unreliable queues.

In Section 3.3 we present further applications and discuss the differences between the stationary distributions of the continuous time process and the embedded Markov chain.

In Section A.1 we provide some useful lemmata on invertibility of matrices which seem to be of interest for their own.

Related work: We have cited literature related to our introductory examples which deal with product form stationary distributions above. Clearly, there are many investigations on queueing systems with unreliable servers without this separability property, for a survey see the introductions in [SD03a] and [Sau06].

In classical Operations Research the fields of queueing theory and inventory theory are almost disjoint areas of research. Recently, research on integrated models has found some interest, a survey is the review in [KLM11].

In a more abstract setting, we can describe our present work as to develop a framework for a birth-death process in a random environment, where the birth-death process' development is interrupted from time to time by some configurations occurring in the environment. On the other side, in our framework the birth-death process influences the development of the environment.

There are many investigations on birth-death processes in random environments, we shall cite only some selected references. Best to our knowledge our results below are complementary to the literature. A stream of research on birth-death processes in a random environment exploits the interaction of birth-death process and environment as the typical structure of a quasi-birth-death process. Such "QBD processes" have two dimensional states, the "level" indicates the population size, while the "phase" represents the environment. For more details see Chapter 6 (Queues in a Random Environment) in [Neu81], and Example C in [Neu89][p. 202, 203].

Related models are investigated in the theory of branching processes in a random environment, see Section 2.9 in [HJV05] for a short review. An early survey with many references to old literature is [Kes80].

Another branch of research is optimization of queues under constraints put on the queue by a randomly changing environment as described e.g. in [HW84].

While the most of the annotated sources are concerned with conventional steady-state analysis, the work [Fal96] is related to ours two-fold: A queue (finite classical loss system) in a random environment shows a product form steady state.

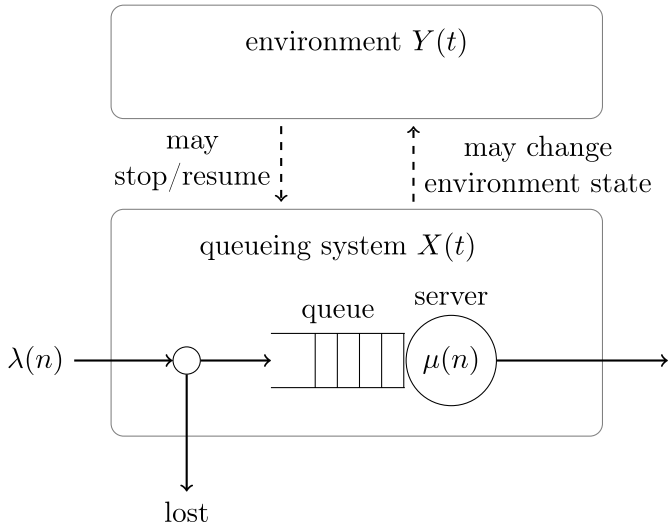


Figure 1.0.1.: Loss system

2. Loss systems in continuous time

Parts of this section are published in [KD15a]. The final publication is available at link.springer.com.

2.1. The exponential model

2.1.1. The $M/M/1/\infty$ model

We consider a two-dimensional process $Z = (X, Y) = ((X(t), Y(t)) : t \in [0, \infty))$ with state space $E = \mathbb{N}_0 \times K$. K is a countable set, the environment space of the process, whereas the queueing state space is $\mathbb{N}_0$.

We assume throughout that $Z = (X, Y)$ is regular and irreducible (unless specified otherwise).

According to our introductory example the environment space of the process is partitioned into disjoint components $K := K_W \uplus K_B$. In the framework of K describing the inventory size K_B describes the status "stock out", in the reliability problem K_B describes the status "server broken down". So accordingly K_W indicates for the inventory that there is stock on hand for production, and "server is up" in the other system.

The general interpretation is that whenever the environment process enters K_B the service process is "**B**LOCKED", and the service is resumed immediately whenever the environment process returns to K_W, the server "**W**ORKS" again.

Whenever the environment process stays in K_B new arrivals are lost.

Obviously, it is natural to assume that the set K_W is not empty, while in certain frameworks K_B may be empty, e.g. no break down of the server in the second introductory example occurs.

The server in the system is a single server under First-Come-First-Served regime (FCFS) with an infinite waiting room.

The arrival stream of customers is Poisson with rate $\lambda(n) > 0$, when there are n customers in the system.

The system develops over time as follows.

1) If the environment at time t is in state $Y(t) = k \in K_W$ and if there are $X(t) = n$ customers in the queue then service is provided to the customer at the head of the queue with rate $\mu(n) > 0$. The queue is organized according First-Come-First-Served regime (FCFS). As soon as his service is finished he leaves the system and the environment changes with probability $R(k, m)$ to state $m \in K$, independent of the history of the system, given k. We consider $R = (R(k, m) : k, m \in K)$ as a stochastic matrix for the environment driven by the departure process.

2) If the environment at time t is in state $Y(t) = k \in K_B$ no service is provided to customers in the queue and arriving customers are lost.

3) Whenever the environment at time t is in state $Y(t) = k \in K$ it changes with rate $v(k, m)$ to state $m \in K$, independent of the history of the system, given k.

Note, that such changes occur independent from the service and arrival process, while the changes of the environment's status under **1)** are coupled with the service process.

From the above description we conclude that the non-negative transition rates of (X, Y) are for $(n, k) \in E$

$$\begin{aligned} q((n,k),(n+1,k)) &= \lambda(n), && k \in K_W, \\ q((n,k),(n-1,m)) &= \mu(n)R(k,m), && k \in K_W, n > 0, \\ q((n,k),(n,m)) &= v(k,m) \in \mathbb{R}_0^+, && k \neq m, \\ q((n,k),(i,m)) &= 0, && \text{otherwise for } (n,k) \neq (i,m). \end{aligned}$$

Note, that the diagonal elements of $Q := (q((n,k),(i,m)) : (n,k),(i,m) \in E)$ are determined by the requirement that row sum is 0.

Remark 2.1.1. It is allowed to have positive diagonal entries $R(k,k)$. R needs not be irreducible, there may exist closed subsets in K.

$v(k,k) = -\sum_{m \in K \setminus \{k\}} v(k,m)$ is required for all $k \in K$ such that

$$V = (v(k,m) : k, m \in K)$$

is a generator matrix.

The Markov process associated with V may have absorbing states, i.e., V then has zero rows.

Remark 2.1.2. We will visualize the dynamics of the environment by environment transition and interaction diagram consisting of colored nodes and colored arrows. We will use the following conventions:

- Square nodes describe the environment states from K.
- Red nodes describe the blocking states, i.e., states from K_B.
- Blue arrows describe possible environment changes independent from the queueing system and correspond to the positive rates of the V matrix.
- Black arrows describe environment changes after services and correspond to the positive entries of the R matrix.

For example the diagrams Figure 2.2.2a on page 44 and Figure 2.2.2b on page 44 describe the behaviors of two different lost-sales systems.

2.1.1.1. Steady-state distribution

Our aim is to compute for an ergodic system explicitly the steady-state and limiting distribution of (X, Y). We can not expect that this will be possible in the general system as described in Section 2.1.1, but fortunately enough we will be able to characterize those systems which admit a product form equilibrium.

Definition 2.1.3. For a loss system $(X(t), Y(t))$ in a state space $E := \mathbb{N}_0 \times K$, whose unique limiting distribution exists, we define

$$\pi := (\pi(n,k) : (n,k) \in E := \mathbb{N}_0 \times K),$$
$$\pi(n,k) := \lim_{t \to \infty} P(X(t) = n, Y(t) = k)$$

and the appropriate marginal limiting distributions

$$\xi := (\xi(n) : n \in \mathbb{N}_0) \qquad \text{with} \qquad \xi(n) := \lim_{t\to\infty} P(X(t) = n),$$

$$\theta := (\theta(k) : k \in K) \qquad \text{with} \qquad \theta(k) := \lim_{t\to\infty} P(Y(t) = k).$$

Remark 2.1.4. It will be convenient to order the state space in the way which is common in matrix analytical investigations, where X is the level process and Y is the phase process. Take on $\mathbb{N}_0$ the natural order and fix a total (linear) order $\preccurlyeq$ on K and define on $E = \mathbb{N}_0 \times K$ the lexicographic order $\prec$ by

$$(m,k),(n,l) \in E \ \text{ then } \ \Big((m,k) \prec (n,l) :\Longleftrightarrow \big[m < n \ \text{ or } \ (m = n \ \text{ and } \ k \preccurlyeq l)\big]\Big). \quad (2.1.1)$$

Some notation which will be used henceforth: I_W is a matrix which has ones on its diagonal elements (k,k) with $k \in K_W$ and 0 otherwise. That is

$$(I_W)_{km} = \delta_{km} 1_{[k \in K_W]}.$$

Theorem 2.1.5. *(i) Denote for $n \in \mathbb{N}_0$*

$$Q_{red}(n) := (q_{red}(n;k,m) : k,m \in K) = \lambda(n) I_W (R - I) + V. \quad (2.1.2)$$

Then the matrices $Q_{red}(n)$ are generator matrices for some homogeneous Markov processes and their entries are

$$\begin{aligned} q_{red}(n;k,m) &= \lambda(n) R(k,m) 1_{[k \in K_W]} + v(k,m), && k \neq m, \\ q_{red}(n;k,k) &= -\Big(1_{[k\in K_W]}\lambda(n)(1 - R(k,k)) + \sum_{m \in K\setminus\{k\}} v(k,m)\Big). \end{aligned} \quad (2.1.3)$$

(ii) For the process (X,Y) the following properties are equivalent:

(a) (X,Y) is ergodic with product form steady-state distribution

$$\pi(n,k) = \underbrace{C^{-1} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)}}_{=:\xi(n)} \theta(k). \quad (2.1.4)$$

(b) The summability condition

$$C := \sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)} < \infty \quad (2.1.5)$$

holds, and the equation

$$\theta \cdot Q_{red}(0) = 0 \quad (2.1.6)$$

admits a unique strictly positive stochastic solution $\theta = (\theta(k) : k \in K)$ which solves also

$$\forall n \in \mathbb{N} : \theta \cdot Q_{red}(n) = 0. \quad (2.1.7)$$

Before proving the theorem some remarks seem to be in order.

Remark 2.1.6. *Irreducibility*: Communication between environment states is governed by the matrices $Q_{\text{red}}(n)$, see (2.1.3). This representation indicates that irreducibility of (X, Y) is the result of an interplay between the one step transition matrix R and the generator matrix V. Neither R nor V is required to be irreducible for its own. An important example where the interplay of two absorbing processes, governed by R and V generate irreducible movements of an inventory, which is represented as an environment, is described in Example 2.2.2 below.

We emphasize that if $V = 0$ then R is irreducible and if R is identity matrix then V must be irreducible.

Furthermore, $R(k, k) > 0$ is allowed, i.e. the environment may not change at a departure instant.

Remark 2.1.7. *Independence*: If $Z = (Z(t) : t \geq 0)$ is stationary, i.e., for $t \geq 0$ holds $P(Z(t) = (n, k)) = \pi(n, k) = \xi(n)\theta(k)$, for all $n, k \in E$, then for any fixed time instant t_0 we have a product form distribution

$$\pi(n, k) = \xi(n) \cdot \theta(k), \qquad (n, k) \in E.$$

Note, that this does not mean that the marginal Processes X and Y are independent. Especially X in general is not Markov for its own, although its stationary distribution is identical to that of a birth-death process with birth rates $\lambda(n)$ and death rates $\mu(n)$, i.e.

$$\xi = \left(\xi(n) := C^{-1} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)} : n \in \mathbb{N}_0 \right). \tag{2.1.8}$$

The observation (2.1.8) is remarkable not only because X in general is not a birth-death process, but also because neither the $\lambda(i)$ are the effective arrival rates (expected number of arrivals per time unit) for queue length i nor the $\mu(i+1)$ the effective service rates (expected maximal number of departures per time unit) for queue length $i+1$. In case of a pure birth-death process without an environment $\lambda(i)$, $\mu(i+1)$ **are** the respective rates.

The conclusion is that both rates are diminished by the influence of the environment by the same portion. It seems to be contra intuition to us that the reduction of $\lambda(i)$ goes in parallel to that of $\mu(i+1)$, while in the running system under queue length i due to Y entering K_B arrivals at rate $\lambda(i)$ are interrupted in parallel to services of rate $\mu(i)$.

The similar problem was noticed already for the case of queueing-inventory processes with state independent service and arrival rates in [SSD$^+$06, Remark 2.8], but in this setting clearly the problem of $\lambda(i)$ versus $\mu(i+1)$ is still hidden.

Remark 2.1.8. *Insensitivity*: The statement (2.1.4) of the theorem can be interpreted as a strong insensitivity property of the system: As long as ergodicity is maintained, the environment can change drastically without changing the steady state of the queue length at any fixed time point. An intuitive interpretation of this result seems to be hard. Especially, this insensitivity can not be a consequence of the form of the control of the inventory or the availability.

Interpretation: We believe that there is an intuitive explanation of a part of the result. The main observation with respect to this is:

Whenever a customer is admitted to the queue, i.e. not lost, he observes the service process as that in a conventional $M/M/1/\infty$ queue with state dependent service and arrival rates, as long as the blocking periods are skipped over.

Saying it the other way round, whenever the environment enters K_B and blocks the service process, the arrival process is blocked as well, i.e. the queueing system is completely frozen and is revived immediately when the environment enters K_W next.

Skipping the problem of i versus $i+1$ discussed in Remark 2.1.7 this observation might explain the form of the marginal stationary distribution of the customer process X, but it does by no means explain the product form of limiting distribution $\pi(n,k) = \xi(n)\theta(k)$.

A similar observation was utilized in [SAH13] in a queueing-inventory system (with state independent service and arrival rates) to construct a related system which obviously has the stationary distribution of X and it is argued that from this follows that the original system shows the same marginal queue length distribution.

The proven insensitivity does not mean, that the time development of the queue length processes with fixed $\lambda(n)$ and $\mu(n)$ is the same under different environment behaviour. This can be seen by considering the stationary sojourn time of admitted customers, which is strongly dependent of the interruption time distributions (= sojourn time distribution of Y in K_B).

Similarly, multidimensional stationary probabilities for $(X(t_1), X(t_2), \ldots, X(t_n))$ will clearly depend on the occurrence frequency of the event $(Y \in K_B)$.

Proof of Theorem 2.1.5 (i). Utilizing the stochastic matrix property $R\mathbf{e} = \mathbf{e}$ and generator property $V\mathbf{e} = 0$ we get:

$$Q_{\text{red}}(n)\mathbf{e} = (\lambda(n)I_W(R-I)+V)\,\mathbf{e} = \lambda(n)I_W\underbrace{(\underbrace{R\mathbf{e}}_{=\mathbf{e}} - \underbrace{I\mathbf{e}}_{=\mathbf{e}})}_{=0} + \underbrace{V\mathbf{e}}_{=0} = 0. \tag{2.1.9}$$

Using the fact that all entries of R and all non-diagonal entries of V are non-negative we see that for all $k \neq m$ it holds

$$(Q_{\text{red}}(n))_{km} = \left(\underbrace{\lambda(n)I_W R}_{\geq 0}\right)_{km} - \underbrace{(\lambda(n)I_W I)_{km}}_{=0} + \underbrace{(V)_{km}}_{\geq 0} \geq 0. \tag{2.1.10}$$

(2.1.9) and (2.1.10) together show that the $Q_{\text{red}}(n)$ are generator matrices. The explicit representation (2.1.3) of the matrix $Q_{\text{red}}(n)$ is calculating directly.

(ii) (b) $\Rightarrow$ (a):

The global balance equations of the Markov process (X,Y) are for $(n,k) \in E$

$$\begin{aligned}
&\pi(n,k)\left(1_{[k\in K_W]}\lambda(n) + \sum_{m\in K\setminus\{k\}} v(k,m) + 1_{[k\in K_W]}1_{[n>0]}\mu(n)\right) \\
&= \pi(n-1,k)1_{[k\in K_W]}1_{[n>0]}\lambda(n-1) + \sum_{m\in K_W} \pi(n+1,m)R(m,k)\mu(n+1) \\
&+ \sum_{m\in K\setminus\{k\}} \pi(n,m)v(m,k).
\end{aligned} \tag{2.1.11}$$

Inserting the proposed product form solution (2.1.4) for $\pi(n,k)$ into the global balance (2.1.11) equations, canceling C^{-1} yields

$$\begin{aligned}
&\theta(k)\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)}\left(1_{[k\in K_W]}\lambda(n)+\sum_{m\in K\setminus\{k\}}v(k,m)+1_{[k\in K_W]}1_{[n>0]}\mu(n)\right)\\
&=\theta(k)\prod_{i=0}^{n-2}\frac{\lambda(i)}{\mu(i+1)}1_{[k\in K_W]}1_{[n>0]}\lambda(n-1)+\sum_{m\in K_W}\theta(m)\prod_{i=0}^{n}\frac{\lambda(i)}{\mu(i+1)}R(m,k)\mu(n+1)\\
&\quad+\sum_{m\in K\setminus\{k\}}\theta(m)\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)}v(m,k),
\end{aligned}\tag{2.1.12}$$

and multiplication with $\prod_{i=0}^{n-1}\left(\frac{\lambda(i)}{\mu(i+1)}\right)^{-1}$ yields

$$\begin{aligned}
&\theta(k)\left(1_{[k\in K_W]}\lambda(n)+\sum_{m\in K\setminus\{k\}}v(k,m)+1_{[k\in K_W]}1_{[n>0]}\mu(n)\right)\\
&=\theta(k)\frac{\mu(n)}{\lambda(n-1)}1_{[k\in K_W]}1_{[n>0]}\lambda(n-1)+\sum_{m\in K_W}\theta(m)\frac{\lambda(n)}{\mu(n+1)}R(m,k)\mu(n+1)\\
&\quad+\sum_{m\in K\setminus\{k\}}\theta(m)v(m,k)
\end{aligned}$$

$$\begin{aligned}
\iff\quad&\theta(k)\left(1_{[k\in K_W]}\lambda(n)+\sum_{m\in K\setminus\{k\}}v(k,m)+1_{[k\in K_W]}1_{[n>0]}\mu(n)\right)\\
&=\theta(k)\mu(n)1_{[k\in K_W]}1_{[n>0]}+\sum_{m\in K_W}\theta(m)\lambda(n)R(m,k)\\
&\quad+\sum_{m\in K\setminus\{k\}}\theta(m)v(m,k)
\end{aligned}$$

$$\begin{aligned}
\iff 0=&-\theta(k)\left(1_{[k\in K_W]}\lambda(n)+\sum_{m\in K\setminus\{k\}}v(k,m)\right)\\
&+\sum_{m\in K_W}\theta(m)\lambda(n)R(m,k)+\sum_{m\in K\setminus\{k\}}\theta(m)v(m,k)
\end{aligned}$$

$$\begin{aligned}
\iff 0\quad=&\;\theta(k)\underbrace{\left\{-\left(1_{[k\in K_W]}\lambda(n)(1-R(k,k))+\sum_{m\in K\setminus\{k\}}v(k,m)\right)\right\}}_{=:q_{\text{red}}(n,k,k)}\\
&+\textstyle\sum_{m\in K\setminus\{k\}}\theta(m)\underbrace{\left(\lambda(n)R(m,k)1_{[m\in K_W]}+v(m,k)\right)}_{=:q_{\text{red}}(n,m,k)}
\end{aligned}\tag{2.1.13}$$

$$\iff \theta Q_{\mathrm{red}}(n) = 0,$$

which is (for all $n \in \mathbb{N}_0$) the condition (2.1.6) and (2.1.7).

By assumption (2.1.6) there exists a stochastic solution to $\theta \cdot Q_{\mathrm{red}}(0) = 0$, which according to requirement (2.1.7) is a solution of $\theta Q_{\mathrm{red}}(n) = 0$.

Taking this θ in (2.1.12) provides a solution of the global balance equations (2.1.11). Thus, with $\pi(n,k) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)} \theta(k)$ we found a positive stochastic solution π of the global balance equations (2.1.11). By assumption the process (X, Y) is regular and irreducible, therefore this π is the unique steady-state distribution of (X, Y).

(ii) (a) $\Rightarrow$ (b):
Because π is stochastic, summability (2.1.5) holds. Insert the stochastic vector of product form (2.1.4) into (2.1.11). As shown in the part (b) $\Rightarrow$ (a) of the proof this leads to (2.1.13) and we have found a solution of (2.1.6) which solves (2.1.7) for all $n \in \mathbb{N}$ as well. $\square$

Proposition 2.1.9. *For a loss system the following three statements are equivalent*

(a) Q is irreducible.

(b) All matrices $Q_{red}(n)$ $n \in \mathbb{N}_0$ are irreducible.

(c) $I_W(R-I)+V$ is irreducible.

Proof. To simplify notation we abbreviate in this proof $M := I_W(R-I)+V$, $M \in \mathbb{R}^{K \times K}$. Consider the directed transition graph of M, with vertices K and edges $\mathcal{E}$ defined by $km \in \mathcal{E} \iff M_{km} > 0$. We will call a finite path from k_0 to k_N in M a sequence of vertices $(k_0, ..., k_N)$, where $M_{k_i,k_{i+1}} > 0$ for all $i \in \{0, \ldots, N-1\}$. The length of this path – number of edges – is N . Similarly we define a path in matrices $Q_{\mathrm{red}}(n)$ and Q.

(b) $\Leftrightarrow$ (c): For any $k, m \in K$ and $\lambda(n) \in \mathbb{R}^+$ it holds

$$(I_W(R-I)+V)_{km} > 0 \iff (\lambda(n) I_W(R-I)+V)_{km} > 0.$$

(a) $\Rightarrow$ (c): Let k and m be arbitrary but fixed states from K with $k \neq m$. Because Q is irreducible, there exists a finite path of length $N \geq 1$

$$\big((n_0, k_0), ..., (n_N, k_N)\big)$$

with $(n_0, k_0) := (1, k)$ and $(n_N, k_N) := (1, m)$, such that $q\left((n_i, k_i), (n_{i+1}, k_{i+1})\right) > 0$ for all $i \in \{0, ..., N-1\}$. In a loss system, there are only three possibilities for *positive* entries $q\left((n_i, k_i), (n_{i+1}, k_{i+1})\right)$:

$$\begin{aligned}
q\left((n_i, k_i), (n_{i+1}, k_{i+1})\right) &= \lambda(n_i) > 0 && \text{in case } n_{i+1} = n_i + 1,\ k_i = k_{i+1}, \\
& && \text{and } k_i \in K_W, \\
q\left((n_i, k_i), (n_{i+1}, k_{i+1})\right) &= \mu(n_i) R(k_i, k_{i+1}) > 0 && \text{in case } n_{i+1} = n_i - 1, \text{ and } k_i \in K_W, \\
q\left((n_i, k_i), (n_{i+1}, k_{i+1})\right) &= v(k_i, k_{i+1}) > 0 && \text{in case } \ n_i = n_{i+1} \text{ and } k_i \neq k_{i+1}.
\end{aligned}$$

Therefore for the sequence $(k_0, ..., k_N)$ with $k_0 = k$ and $k_N = m$ and $M := I_W(R-I)+V$ it hold for all $k_i \neq k_{i+1}$

$$M_{k_i,k_{i+1}} = 1_{[k_i \in K_W]} R(k_i, k_{i+1}) + v(k_i, k_{i+1}) \geq \begin{cases} R(k_i, k_{i+1}) > 0 & \text{if } n_{i+1} = n_i - 1, \\ v(k_i, k_{i+1}) > 0 & \text{if } n_{i+1} = n_i. \end{cases}$$

Removing all k_i with $k_i = k_{i+1}$ from the sequence $(k_0, ..., k_N)$ we get a finite path $(k_0^{'}, ..., k_{N'}^{'})$ in M from k to m.

(c) $\Rightarrow$ (a): We have to prove, that for any $n, \ell \in \mathbb{N}_0$ and any $k, m \in K$ with $(n, k) \neq (\ell, m)$ there exists a finite path $\big((n_0, k_0), \ldots, (n_N, k_N)\big)$ from (n, k) to (ℓ, m) in Q i.e $(n_0, k_0) = (n, k)$, $(n_N, k_N) = (\ell, m)$ and $q\left((n_i, k_i), (n_{i+1}, k_{i+1})\right) > 0$ for all $i \in \{0, ..., N-1\}$. Before we will prove this statement for any states $(n, k) \neq (\ell, m)$ we show how to construct finite paths in E for particular states.

(i) For any k and any m with $k \neq m$ and $M_{km} > 0$ and any $n \in \mathbb{N}_0$ it is possible to construct a path of maximal length 2 from (n, k) to (n, m) in Q: If $v(k, m) > 0$ take the path $\big((n, k), (n, m)\big)$ since it holds $q((n, k), (n, m)) = v(k, m) > 0$. If $v(k, m) = 0$ and $M_{km} > 0$ then it must hold $(I_W(R - I))_{km} > 0$. This implies $k \in K_W$ and $R(k, m) > 0$. In this case, use the path $\big((n, k), (n + 1, k), (n, m)\big)$ in E. For this path it holds $q((n, k), (n + 1, k)) = \lambda(n) > 0$ and $q((n + 1, k), (n, m)) = 1_{[k \in K_W]}\mu(n + 1)R(k, m) > 0$.

(ii) For any $k \in K$ and $m \in K$ with $k \neq m$ and any $n \in \mathbb{N}_0$ there is a finite path $\big((n_0, k_0), \ldots, (n_N, k_N)\big)$ from (n, k) to (n, m) in Q: Because M is irreducible there is a finite path $(k_0, \ldots, k_s)$ from $k = k_0$ to $m = k_s$. A path in M means that for each edge (k_i, k_{i+1}) it holds $M_{k_i, k_{i+1}} > 0$. Using construction of path in Q described in (i) for each edge (k_i, k_{i+1}), we construct a path in Q from (n, k_0) to (n, k_s). Its size is less or equal $2s$.

(iii) For any $k \in K$ and any $n \in \mathbb{N}_0$ there exists some state $m \in K$ with a finite path from (n, k) to $(n + 1, m)$ in Q: If $k \in K_W$ then $m = k$ and the path in Q is $\big((n, k), (n + 1, k)\big)$. It holds $q((n, k), (n + 1, m)) = \lambda(n) > 0$. If $k \in K_B$ take any $m \in K_W$ and construct a path $\big((n_0, k_0), \ldots, (n_N, k_N)\big)$ from (n, k) to (n, m) using method (ii). Finally paste $(n + 1, m)$ to the path.

(iv) For any $k \in K$ and any $n \in \mathbb{N}$ there exists some state $m \in K$ with a finite path (n, k) to $(n - 1, m)$ in Q: If $k \in K_W$ then take any m such that $R(k, m) > 0$ and the path in Q is $\big((n, k), (n - 1, m)\big)$. It holds $q((n, k), (n - 1, m)) = 1_{[k \in K_W]}\mu(n)R(k, m) > 0$. If $k \in K_B$ take any $m' \in K_W$ and $m \in K$ such that $R_{(m', m)} > 0$. Construct a path $(n_1, k_1), \ldots, (n_N, k_N)$ from (n, k) to (n, m') using method (ii). Finally add $(n - 1, m)$ to the path.

Let $n, \ell \in \mathbb{N}_0$ and $k, m \in K$ arbitrary states with $(n, k) \neq (\ell, m)$. If $n = \ell$ we can use the method (ii) to construct a path from (n, k) to (ℓ, m). If $n < \ell$ then we use the method (iii) $\ell - n$ times to construct a path to some element (ℓ, m') and then append a path from (ℓ, m') to (ℓ, m) constructed by the method (ii) to it. Similarly, if $n > \ell$ then we use the method (iv) $n - \ell$ times to construct a path to some element (ℓ, m') and then append a path from (ℓ, m') to (ℓ, m) constructed by the method (ii) to it. $\square$

Most of the examples in this thesis deal with loss systems with constant arrival rate λ. To simplify analysis for these systems we summarize the results for this special case of Theorem 2.1.5 and some results Proposition 2.1.9 in the next corollary.

Corollary 2.1.10. [1] *Consider the framework of Theorem 2.1.5 with a Poisson-λ arrival stream (which is interrupted when the environment process stays in K_B) i.e. $\lambda(n) = \lambda$ for all $n \in \mathbb{N}_0$.*

(i) Denote

$$Q_{red} := \lambda I_W(R - I) + V. \tag{2.1.14}$$

Then the matrix Q_{red} is an irreducible generator matrix for some homogeneous Markov processes

(ii) For the process (X, Y) the following properties are equivalent:

(a) (X, Y) is ergodic with product form steady-state distribution

$$\pi(n, k) = \xi(n)\theta(k) \quad (n, k) \in E, \tag{2.1.15}$$

with

$$\xi(n) := C^{-1}\frac{\lambda^n}{\prod_{i=0}^{n-1}\mu(i+1)}. \tag{2.1.16}$$

(b) The summability condition

$$C := \sum_{n=0}^{\infty}\frac{\lambda^n}{\prod_{i=0}^{n-1}\mu(i+1)} < \infty \tag{2.1.17}$$

holds, and the equation

$$\theta\underbrace{\left(\lambda I_W\left(R - I\right) + V\right)}_{=Q_{red}} = 0. \tag{2.1.18}$$

admits a unique strictly positive stochastic solution.

Proof. (i) Q_{red} is a generator according to Theorem 2.1.5 (i). By assumption Q is irreducible, from Proposition 2.1.9 follows $Q_{\text{red}}(0) = Q_{\text{red}}$ is irreducible.

(ii) follows from Theorem 2.1.5 (ii). Because of $\lambda(n) = \lambda$ for all n holds $Q_{\text{red}}(0) = Q_{\text{red}}(n)$ and the condition (2.1.7) is trivially valid. Equation (2.1.18) is the condition (2.1.6) expressed via matrix representation (2.1.2) of $Q_{\text{red}}(0)$. □

Lemma 2.1.11. *For an ergodic loss system $(X(t), Y(t))$ with finite set K it holds for all $n \in \mathbb{N}_0$*

$$\lim_{t\to\infty} P(X(t) = n | Y(t) \in K_W) = \frac{\sum_{k\in K_W}\pi(n,k)}{\sum_{n\in\mathbb{N}_0}\sum_{k\in K_W}\pi(n,k)} = C^{-1}\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)}$$

and

$$C := \sum_{n=0}^{\infty}\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)} < \infty.$$

[1] The implication (b) ⇒(a) is published as [KD15a, Theorem 1].

Proof. Summation of the global balance equations over $k \in K$ leads to

$$\sum_{k\in K}\pi(n,k)\left(1_{[k\in K_W]}\lambda(n)+\underbrace{\sum_{m\in K\setminus\{k\}}v(k,m)}_{-v(k,k)}+1_{[k\in K_W]}1_{[n>0]}\mu(n)\right)$$
$$=\sum_{k\in K}\pi(n-1,k)1_{[k\in K_W]}1_{[n>0]}\lambda(n-1)+\sum_{k\in K}\sum_{m\in K_W}\pi(n+1,m)R(m,k)\mu(n+1)$$
$$+\sum_{k\in K}\sum_{m\in K\setminus\{k\}}\pi(n,m)v(m,k)$$

$$\iff \sum_{k\in K_W}\pi(n,k)\left(\lambda(n)+1_{[n>0]}\mu(n)\right)$$
$$=\sum_{k\in K_W}\pi(n-1,k)1_{[n>0]}\lambda(n-1)+\sum_{m\in K_W}\pi(n+1,m)\underbrace{\sum_{k\in K}R(m,k)}_{=1}\mu(n+1)$$
$$+\sum_{m\in K}\pi(n,m)\underbrace{\sum_{k\in K}v(m,k)}_{=0}$$

$$\iff \sum_{k\in K_W}\pi(n,k)\left(\lambda(n)+1_{[n>0]}\mu(n)\right)$$
$$=\sum_{k\in K_W}\pi(n-1,k)1_{[n>0]}\lambda(n-1)+\sum_{m\in K_W}\pi(n+1,m)\mu(n+1).$$

We decide the last equation by $\sum_{\ell\in\mathbb{N}_0}\sum_{k\in K_W}\pi(\ell,k)$ and get with

$$\xi_W(n):=\frac{\sum_{k\in K_W}\pi(n,k)}{\sum_{\ell\in\mathbb{N}_0}\sum_{k\in K_W}\pi(\ell,k)}.$$

$$\iff \xi_W(n)\left(\lambda(n)+1_{[n>0]}\mu(n)\right)=\xi_W(n-1)1_{[n>0]}\lambda(n-1)+\xi_W(n+1)\mu(n+1). \quad (2.1.19)$$

Equation (2.1.19) is the steady-state equation for a some irreducible birth-and-death process with positive birth rates $\lambda(n)$, positive death rates $\mu(n)$, and positive normalized solution $\xi_W(n)$.

It therefore holds $\xi_W(n)=\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)}\xi_W(0)$. Therefore $C=\sum_{n=0}^{\infty}\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)}=\sum_{n=0}^{\infty}\frac{\xi_W(n)}{\xi_W(0)}=\frac{1}{\xi_W(0)}<\infty$. □

Remark 2.1.12. In Lemma 2.1.11 we required the set K to be finite in order to keep the proof simple. The problem with an infinite set K lies in the expression

$$\sum_{k\in K}\pi(n,k)\underbrace{\sum_{m\in K\setminus\{k\}}v(k,m)}_{-v(k,k)}$$

on the LHS of the very first equation in the proof. We need to prove that this expression is finite. In this case all the sums in equations are absolute convergent series and we can use the same proof as for finite K. For example this expression is bounded if V is uniformizable, i.e. if it holds

$$\sup_{k\in K} |v(k,k)| < \infty. \tag{2.1.20}$$

Therefore the statements Lemma 2.1.11 are also true for infinite set K if (2.1.20) holds.

The following corollary states that, for ergodic loss systems with constant arrival rate λ and finite set K, the product-form distribution (2.1.15) with particular ξ and θ follows directly. The conditions (2.1.17), (2.1.16) and (2.1.18) are in this case redundant.

Corollary 2.1.13. *Given an ergodic loss system with constant arrival rate λ and finite set K then it holds*

$$\pi(n,k) = \xi(n)\theta(k)$$

with ξ defined in (2.1.16) *and with θ the unique positive stochastic solution of the equation* (2.1.18)*.*

Proof. From Lemma 2.1.11 we have summability condition (2.1.17). According to Corollary 2.1.10 (i) $Q_{\text{red}} \in \mathbb{R}^{K\times K}$ is an irreducible generator matrix. K is finite therefore there exists a unique positive stochastic solution of the equation (2.1.18). The particular product form of π in (2.1.15) with (2.1.16) and (2.1.18) follows from Corollary 2.1.10 (ii). □

Corollary 2.1.14. *Given an ergodic loss system with finite set K with product-form steady-state distribution, then it holds*

$$\pi(n,k) = \xi(n)\theta(k)$$

with ξ defined in (2.1.4) *and θ is unique positive stochastic solution of the equations* (2.1.7)*.*

Proof. If the steady-state distribution has a product form then it holds for any $n \in \mathbb{N}_0$

$$\xi(n) = \lim_{t\to\infty} P(X(t)=n) = \lim_{t\to\infty} P(X(t)=n|Y(t)=k) \overset{\text{Lemma 2.1.11}}{=} C^{-1}\prod_{i=0}^{n-1}\frac{\lambda(i)}{\mu(i+1)}.$$

Applying Theorem 2.1.5 (ii) proves that θ is unique positive stochastic solution of the equations (2.1.7). □

The next examples comment on different forms of establishing product form equilibrium which may arise in the realm of Theorem 2.1.5.

Example 2.1.15. [2]There exist non-trivial ergodic loss systems with non constant (i.e., state dependent) arrival rates $\lambda(n)$ in a random environment which have a product form steady-state distribution. This is verified by the following example.

In any ergodic loss system with finite environment state space K and $V = 0$ it holds for the stochastic solutions θ_n of the equations (2.1.6) and (2.1.7)

$$\theta_n\left(\lambda(n)I_W(R-I)+V\right) = 0 \Longleftrightarrow \theta_n I_W(R-I) = 0.$$

[2]This is a more detailed version of [KD15a, Example 1] but with a finite environment state space.

Because the process is ergodic and $V = 0$, R must be irreducible and $I_W = I$. The stochastic vectors θ_n are then the unique positive stochastic solution of the equations

$$\theta_n(R - I) = 0 \qquad \forall n \in \mathbb{N}_0.$$

Therefore $\theta_n = \theta_0 =: \theta$ for all n. The condition $C := \sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)} < \infty$ follows from Lemma 2.1.11.

Example 2.1.16. [3]There exist a non-trivial ergodic loss system in random environments which has a product form steady-state distribution if and only if the arrival rates are independent of the queue lengths, i.e. $\lambda(n) \equiv \lambda$. This is verified by the following example (which describes a queueing-inventory system under (r, S) policy with $(r = 1, S = 2)$, as will be seen in Section 2.2.1, Definition 2.2.1). We have an environment

$$K = \{0, 1, 2\}, \quad \text{with blocking set} \quad K_B = \{0\},$$

stochastic matrix R and and the generator matrix V given as

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} -\nu & 0 & \nu \\ 0 & -\nu & \nu \\ 0 & 0 & 0 \end{pmatrix}.$$

If this system has a product form steady-state distribution then according to Corollary 2.1.14 this distribution has the form (2.1.4) and we can apply (2.1.7) here.

Clearly, if $\lambda(n) \equiv \lambda$ are equal, the equations

$$\theta \cdot Q_{\text{red}}(n) = 0, \quad n \in \mathbb{N}_0, \tag{2.1.21}$$

with

$$Q_{\text{red}}(n) = \begin{pmatrix} -\nu & 0 & \nu \\ \lambda(n) & -(\lambda(n) + \nu) & \nu \\ 0 & \lambda(n) & -\lambda(n) \end{pmatrix}$$

have a common stochastic solution.

On the other hand, the solutions of (2.1.21) are

$$\theta_n = (\theta_n(0), \theta_n(1), \theta_n(2)) = C_n^{-1} \left(\frac{\lambda(n)}{\nu}, 1, \frac{\lambda(n) + \nu}{\lambda(n)} \right), \quad n \in \mathbb{N}_0\,. \tag{2.1.22}$$

We conclude

$$\forall n \in \mathbb{N}_0 : \theta_n = \theta_{n+1} \Longrightarrow \frac{\theta_n(0)}{\theta_n(1)} = \frac{\theta_{n+1}(0)}{\theta_{n+1}(1)} \Longleftrightarrow \lambda(n) = \lambda(n+1).$$

Remark 2.1.17. In Section 2.1.2 we will show in the course of proving a companion of Theorem 2.1.5 for loss systems with finite waiting room that more restrictive conditions on the environment are needed. It turns out that the construction in the proof of the Theorem 2.1.18 will provide us with more general constructions for examples as those given here, see Remark 2.1.22 below.

[3]This is a [KD15b, Example 2], but with a more general statement. In this thesis the statement is about any product form instead of a particular product form in [KD15b, Example 2].

2.1.1.2. Limiting case

Consider a process Z as described in Section 2.1.1 and set the service times to zero. That means the service rate becomes infinite. Obviously the system will never hold customers in the queue once it became empty. Therefore, for the long-term-behavior analysis we can restrict the state space to $E_{\mu=\infty} := \{0\} \times K$, and consider a non-explosive irreducible process $Z_{\mu=\infty} = (X, Y) = ((X(t), Y(t)) : t \in [0, \infty))$ on $E_{\mu=\infty}$ with generator $Q_{\mu=\infty} := (q_{\mu=\infty}((0, k), (0, m)) : (0, k), (0, m) \in E_{\mu=\infty})$

$$
\begin{aligned}
q_{\mu=\infty}((0,k),(0,m)) &= \lambda(0)R(k,m), && k \in K_W, n > 0, && (2.1.23)\\
q_{\mu=\infty}((0,k),(0,m)) &= v(k,m) \in \mathbb{R}_0^+, && k \neq m, && (2.1.24)\\
q_{\mu=\infty}((0,k),(0,m)) &= 0, && \text{otherwise for } (0,k) \neq (0,m).
\end{aligned}
$$

(2.1.23) means that, as soon a customer arrives to the empty and not blocked queue with rate $\lambda(0)$ he is immediately served and the environment changes from k to m with probability $R(k, m)$. (2.1.24) means that the environment changes its state from k to m with rate $v(k, m)$.

The process $Z_{\mu=\infty} = (X, Y) = ((X(t), Y(t)) : t \in [0, \infty))$ is ergodic if and only if there exists a unique positive steady-state distribution $\pi_{\mu=\infty}$ with

$$\pi_{\mu=\infty} Q_{\mu=\infty} = 0.$$

The matrix $Q_{\mu=\infty}$ is related to matrix $Q_{\text{red}}(0) := (\lambda(0)I_W(R - I) + V)$ defined in Theorem 2.1.5, namely every entry $Q_{\mu=\infty}\big((0, m), (0, k)\big)$ equals the entry $Q_{\text{red}}(0)(k, m)$. According to Theorem 2.1.5, in an ergodic system the matrix $Q_{\text{red}}(0) := (\lambda(0)I_W(R - I) + V)$ has a unique strictly positive steady-state solution θ of the equation $\theta Q_{\text{red}}(0) = 0$. Therefore for any ergodic loss system Z with finite service rates and R, V, K, K_W as $Z_{\mu=\infty}$ which has product-form steady-state distribution $\pi(n, k) = \xi(n)\theta(k)$, and the same parameters $\lambda(0)$, it holds

$$\theta(k) = \pi_{\mu=\infty}(0, k) \qquad \forall k \in K.$$

Thus, for loss systems with finite service rates defined in Section 2.1.1, a system with zero service times can be used to discover the marginal distribution θ. In [KN13] and [KML13] the authors use this fact to obtain the marginal steady-state distribution of an inventory.

An interesting property of the distribution $\pi_{\mu=\infty}$ is that it corresponds to a limiting case of steady-state distribution of a loss system with positive service time when $\mu(1) \to \infty$. We demonstrate this by looking at the limit of queuing distribution $\xi(n)$ as function of $\mu(1)$

$$\lim_{\mu(1)\to\infty} \xi(n) = \lim_{\mu(1)\to\infty} \frac{\prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)}}{\sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)}} = \begin{cases} 1 & \text{for } n = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.25}$$

Consequently it holds

$$\pi_{\mu=\infty}(0, k) = \theta(k) = \lim_{\mu(1)\to\infty} \xi(0)\theta(k) = \lim_{\mu(1)\to\infty} \pi(0, k). \tag{2.1.26}$$

Remark. Let Z be an ergodic loss system with finite service rates $\boldsymbol{\mu} := (\mu(n) : n \in \mathbb{N})$ as defined in Section 2.1.1 and with steady-state distribution π. Let $(\boldsymbol{\mu}_u)_{u \in \mathbb{N}_0}$ be a sequence of service rate vectors with elements from $(\mathbb{R}^+)^{\mathbb{N}}$, e.g. $\boldsymbol{\mu}_u = (\mu_u \in \mathbb{R}^+ : u \in \mathbb{N}_0)$, such that $\boldsymbol{\mu}_0 = \boldsymbol{\mu}$ and

$$\mu_u(n) = \mu(n), \quad \forall u, \forall n > 1, \qquad \text{and} \qquad \mu_u(1) \nearrow \infty \text{ for } u \to \infty.$$

We construct a sequence of processes Z_u from $Z =: Z_0$, by changing service rate parameters from $\boldsymbol{\mu}$ to $\boldsymbol{\mu}_u$. The resulting processes Z_u are ergodic, each with a unique positive steady-state distribution π_u. There exists a pointwise limiting measure of the sequence $(\pi_u : u \in \mathbb{N}_0)$ which we denote by

$$\pi_\infty := \lim_{u \to \infty} \pi_u.$$

Using the same calculation as we performed for $\pi_{\mu=\infty}$ in (2.1.25) and (2.1.26) we see that

$$\pi_\infty(n,k) = \lim_{u \to \infty} \pi_u(n,k) = \begin{cases} \theta(k), & \text{for } n = 0\,, \forall k, \\ 0, & \text{for } \forall n > 0\,, \forall k. \end{cases}$$

That means the discrete measure π_∞ is stochastic with support $\{0\} \times K$ and is equal $\pi_{\mu=\infty}$ on $\{0\} \times K$. From point-wise convergence $\lim_{u \to \infty} \pi_u(n,k) = \pi_\infty(n,k)$ we conclude weak convergence $\pi_u \longrightarrow \pi_\infty$.

2.1.2. Finite capacity loss systems

In this section we study the systems from Section 2.1.1 under the additional restriction that the capacity of the waiting room is finite. That is, we now consider loss systems in the traditional sense with the additional feature of losses due to the environment's restrictions on customers' admission and service

Recall, that for the pure exponential single server queueing system with state dependent rates and $N \geq 0$ waiting places the state space is $E = \{0, 1, \ldots, N, N+1\}$ and the queueing process X is ergodic with stationary distribution $\pi = (\pi(n) : n \in E)$ of the form

$$\pi(n) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)}, \quad n \in E. \tag{2.1.27}$$

If the queueing system with infinite waiting room and the same rates $\lambda(i)$, $\mu(i)$ is ergodic, the stationary distribution π in (2.1.27) is simply obtained by conditioning the stationary distribution of this infinite system onto E. (Note, that ergodicity in the finite waiting room case is granted by free, without referring to the infinite system.)

We will show, that a similar construction by conditioning is in general not possible for the loss system in a random environment. The structure of the environment process will play a crucial role for enabling such a conditioning procedure.

We take the interaction between the queue length process X and the environment process Y of the same form as in Section 2.1.1, with R and V of the same form, and $\lambda(i) > 0$ for $i = 0, \ldots, N$, and $\mu(i) > 0$ for $i = 1, \ldots, N+1$. The state space is

$E := \{0, \dots, N+1\} \times K$. The non-negative transition rates of (X, Y) are for $(n, k) \in E$

$$\begin{aligned} q((n,k),(n+1,k)) &= \lambda(n), && k \in K_W, n < N+1, \\ q((n,k),(n-1,m)) &= \mu(n)R(k,m), && k \in K_W, n > 0, \\ q((n,k),(n,m)) &= v(k,m) \in \mathbb{R}_0^+, && k \neq m, \\ q((n,k),(i,m)) &= 0, && \text{otherwise for } (n,k) \neq (i,m) \in E. \end{aligned}$$

The first step of the investigation is nevertheless completely parallel to Theorem 2.1.5.

Theorem 2.1.18. [4]

(i) Denote for $n \in \{0, \dots, N+1\}$

$$\begin{aligned} q_{red}(n,k,m) &= \lambda(n)R(k,m)1_{[k \in K_W]} \cdot 1_{[n \in \{0,\dots,N\}]} + v(k,m), \qquad k \neq m \\ q_{red}(n,k,k) &= -\Big(1_{[k \in K_W]} \cdot 1_{[n \in \{0,\dots,N\}]} \lambda(n)(1 - R(k,k)) + \sum_{m \in K \setminus \{k\}} v(k,m)\Big) \end{aligned} \tag{2.1.28}$$

and

$$Q_{red}(n) = (q_{red}(n,k,m) : k, m \in K).$$

Then the matrices $Q_{red}(n)$ are generator matrices for some homogeneous Markov processes.

(ii) If the process (X, Y) is ergodic denote its unique steady-state distribution by

$$\pi = (\pi(n,k) : (n,k) \in E := \{0, \dots, N+1\} \times K).$$

Then the following three properties are equivalent:

(a) (X, Y) is ergodic on E with product form steady-state distribution

$$\pi(n,k) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)} \theta(k) \quad n \in \{0, \dots, N+1\}, k \in K. \tag{2.1.29}$$

(b) The equation

$$\theta \cdot Q_{red}(0) = 0 \tag{2.1.30}$$

admits a strict positive stochastic solution $\theta = (\theta(k) : k \in K)$ which solves also

$$\forall n \in \{0, \dots, N+1\} : \theta \cdot Q_{red}(n) = 0. \tag{2.1.31}$$

(c) The equation

$$\eta \cdot V = 0 \tag{2.1.32}$$

admits a strict positive stochastic solution.
The set $K_W \subset K$ is a closed set for the Markov chain on state space K with transition matrix R, i.e.,

$$\forall k \in K_W : \quad \sum_{m \in K_W} R(k,m) = 1,$$

[4] The implication (iii)⇒(i) is published as [KD15a, Theorem 2].

and the restriction $\eta^{(W)} := (\eta(m) : m \in K_W)$ of η to K_W solves the equation

$$\eta^{(W)} = \eta^{(W)} \cdot R^{(W)}, \tag{2.1.33}$$

where

$$R^{(W)} := (R(k,m) : k, m \in K_W) \tag{2.1.34}$$

is the restriction of R to K_W.

Proof. The proof of (i) is similar to that of Theorem 2.1.5 (i), and in (ii) the equivalence of (a) and (b) is proven in almost identical way as that of Theorem 2.1.5 (ii) (with the obvious slight changes due to having the X-component finite) and are therefore omitted.

We next show

(ii) (a) $\Rightarrow$ (c):

The global balance equations of the Markov process (X, Y) are for $(n, k) \in E$

$$\begin{aligned}&\pi(n,k)\left(1_{[k\in K_W]}\cdot 1_{[n\in\{0,\dots,N\}]}\lambda(n) + \sum_{m\in K\setminus\{k\}} v(k,m) + 1_{[k\in K_W]}1_{[n>0]}\mu(n)\right)\\ &= \pi(n-1,k)1_{[k\in K_W]}1_{[n>0]}\lambda(n-1) + \sum_{m\in K_W}\pi(n+1,m)R(m,k)\mu(n+1)\cdot 1_{[n\in\{0,\dots,N\}]}\\ &\quad + \sum_{m\in K\setminus\{k\}}\pi(n,m)v(m,k).\end{aligned} \tag{2.1.35}$$

Inserting the proposed product form solution (2.1.29) for $\pi(n, k)$ into the global balance equations (2.1.35) and proceeding in the same way as in the proof of Theorem 2.1.5 yields

$$\begin{aligned}0 = &-\theta(k)\left(1_{[k\in K_W]}\cdot 1_{[n\in\{0,\dots,N\}]}\lambda(n) + \sum_{m\in K\setminus\{k\}} v(k,m)\right)\\ &+ \sum_{m\in K_W}\theta(m)\lambda(n)R(m,k)\cdot 1_{[n\in\{0,\dots,N\}]} + \sum_{m\in K\setminus\{k\}}\theta(m)v(m,k).\end{aligned} \tag{2.1.36}$$

For $n \to N + 1$ (2.1.36) turns to

$$\theta(k)\left(\sum_{m\in K\setminus\{k\}} v(k,m)\right) = \sum_{m\in K\setminus\{k\}}\theta(m)v(m,k), \tag{2.1.37}$$

which verifies (2.1.32) with $\eta := \theta$.

For $n < N + 1$ (2.1.36) turns to

$$\theta(k)1_{[k\in K_W]}\lambda(n) + \underbrace{\theta(k)\sum_{m\in K\setminus\{k\}} v(k,m)}_{(*)} = \sum_{m\in K_W}\theta(m)\lambda(n)R(m,k) + \underbrace{\sum_{m\in K\setminus\{k\}}\theta(m)v(m,k)}_{(**)}$$

where from (2.1.37) the expressions $(**)$ and $(*)$ cancel and we arrive at

$$\theta(k)1_{[k\in K_W]}\lambda(n) = \sum_{m\in K_W}\theta(m)\lambda(n)R(m,k). \tag{2.1.38}$$

Because (X, Y) is ergodic, θ is strictly positive, and we conclude (set $k \in K_B$ in (2.1.38) which makes the left side zero)

$$R(m, k) = 0, \quad \forall m \in K_W,\ k \in K_B,$$

which shows that K_W is a closed set for the Markov chain governed by R.

Now set $k \in K_W$ in (2.1.38) which makes the left side strictly positive and realize that this after canceling $\lambda(n)$ is exactly (2.1.33).

This part of the proof is finished.

(ii) (c)$\Rightarrow$ (b):

For proving the reversed direction we reconsider the previous part (a) $\Rightarrow$ (c) of the proof: The strict positive stochastic solution of

$$\eta \cdot V = 0, \tag{2.1.39}$$

which is given by assumption (2.1.32), yields the required solution for $n \to N+1$ of

$$\theta \cdot Q_{\text{red}}(N+1) = 0.$$

If $K_W \subset K$ is a closed set for the Markov chain on state space K with transition matrix R we obtain

$$R(m, k) = 0, \quad \forall m \in K_W,\ k \in K_B,$$

and therefore for all $n \in \{0, 1, \ldots, N\}$

$$\theta \cdot Q_{\text{red}}(n) = 0$$

reduces for $k \in K_B$ to the respective expression in

$$\eta \cdot V = 0.$$

It remains for all $n \in \{0, 1, \ldots, N\}$ and for $k \in K_W$ to show that for $k \in K_W$ the respective expression in

$$\theta \cdot Q_{\text{red}}(n) = 0$$

is valid. This follows by considering

$$\eta(k)1_{[k \in K_W]}\lambda(n) + \underbrace{\eta(k) \sum_{m \in K \setminus \{k\}} v(k, m)}_{(*)} = \sum_{m \in K_W} \eta(m)\lambda(n)R(m, k) + \underbrace{\sum_{m \in K \setminus \{k\}} \eta(m)v(m, k)}_{(**)}$$

and remembering that the expressions $(**)$ and $(*)$ cancel. The residual terms are equal by the assumption (2.1.33).

This finishes the proof. $\square$

The interesting insight is that from the existence of the product form steady-state distribution π on $E = \{0, \ldots, N+1\} \times K$ implicitly restrictions on the form of the movements of the environment emerge which are not necessary in the case of infinite waiting rooms. (As indicated above, such restrictions are not necessary too in the pure queueing system framework.)

The proof of Theorem 2.1.18 has brought out the following additional, somewhat surprising, insensitivity property.

Corollary 2.1.19. [5] *Whenever (X, Y) is ergodic with product form steady-state distribution*

$$\pi(n,k) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)} \theta(k) \quad n \in \{0, \dots, N+1\}, k \in K$$

for some (positive) parameter setting $(\lambda(i) : i = 0, 1, \dots, N)$, $(\mu(i) : i = 1, \dots, N+1)$ with an environment characterized by (K, K_B, V, R), then for this same environment (X, Y) is ergodic with product form steady-state distribution with the same θ for any (positive) parameter setting for the arrival and service rates.

Proof. This becomes obvious at the step where we arrived at (2.1.38) and we see that the specific shape of the sequence of the $\lambda(i)$ do not matter. The specific $\mu(i)$ are canceled in the early steps of the proof already. □

Example 2.1.20. We describe a class of examples of environments which guarantee that the conditions of Theorem 2.1.18 are fulfilled. The construction is in three steps.

Take for V a generator of an irreducible Markov process on K with stationary distribution θ, which fulfills for all $k \in K_W$ the *partial balance condition*

$$\theta(k) \sum_{m \in K_W} v(k,m) = \sum_{m \in K_W} \theta(m) v(m,k) \tag{2.1.40}$$

and $\sup(-v(k,k) : k \in K_W) < \infty$.

Denote by $V^{(W)}$ the restriction of V onto K_W which has stationary distribution $\theta^{(W)} := (\theta(k)/(\sum_{m \in K_W} \theta(m)) : k \in K_W)$, see [Kel79, Exercise 1.6.2, p. 27].

Take for $R^{(W)}$ (see (2.1.34)) a uniformization chain of $V^{(W)}$, see [Kei79, Chapter 2, Section 2.1], e.g., (with I the identity matrix on K_W)

$$R^{(W)} := I + \sup(-v(k,k) : k \in K_W)^{-1} V^{(W)},$$

which is stochastic and has equilibrium distribution $\theta^{(W)} := (\theta(k)/(\sum_{m \in K_W} \theta(m)) : k \in K_W)$ as well.

$(R(k,m) : k \in K_B, m \in K)$ can be arbitrarily selected, e.g. the identity matrix on K_B.

This construction ensures that the restriction $\eta^{(W)} := (\eta(m) : m \in K_W)$ of η to K_W solves the equation (2.1.33)

$$\eta^{(W)} = \eta^{(W)} \cdot R^{(W)}.$$

Remark 2.1.21. The construction in Example 2.1.20 may seem to produce a narrow class of examples, but this is not so: All reversible V fulfill the partial balance condition (2.1.40).

Remark 2.1.22. The construction above produces another example contributing to the discussion at the end of Section 2.1.1.1 on the question which particular product forms can occur, and which form of the environment and the arrival and service rate patterns may interact to result in product form equilibrium for loss systems with infinite waiting room. We only have to notice that the equations for $n < N + 1$ are exactly those which occur for all $n \in \mathbb{N}_0$ in the setting of Theorem 2.1.5.

The cautious reader will already have noticed that the conditions in (ii) **(iii)** of Theorem 2.1.18 provide a similar more abstract example for the discussion on Theorem 2.1.5 at the end of Section 2.1.1.1.

[5] published as [KD15a, Corollary 1 in Section 3]

Remark 2.1.23. We should point out that in [SSD+06, Section 6] queueing-inventory models with finite waiting room are investigated with a resulting "quasi product form" steady-state distribution. The respective theorems there do not fit into the realm of Theorem 2.1.18 because the state space is **not** a product space as in Theorem 2.1.18, where we have irreducibility on $E = \{0, 1, \dots, N, N+1\}$.

The difference is that in [SSD+06, Section 6] the element (in notation of the present paper) $(N+1, 0)$ is not a feasible state.

The results there can be considered as a truncation property of the equilibrium of the system with infinite waiting room onto the feasible state space under restriction to finite queues.

2.1.3. Loss systems and matrix-geometric methods

In this section we will analyze loss systems from the point of view of matrix-geometric methods and – related to them – the operator-analytic methods. The matrix-geometric method is a universal and powerful tool to analyze quasi birth-death (QBD) processes. They were popularized in 1970s and are of increasing importance in applied probability nowadays. Loss systems are special cases of QBD processes if the queue length is called *level* and the environment states are called *phase*.

We start with loss system with constant arrival rates λ and constant service rate μ. The resulting loss system is then a level-independent QBD process. We also restrict the number of environment states to be finite, in order to apply the results of [Neu81] and [LR99] directly.

Definition 2.1.24 (Level-independent QBD with finite number of phases)**.** Let $Z = (X(t), Y(t) : t \in \mathbb{R}_0)$ be a level-independent quasi birth-death process on a state space $\mathbb{N}_0 \times K$, where $X(t) \in \mathbb{N}_0$ describes a level and $Y(t) \in K$ describes a phase. The set K is finite. The generator Q of Z has tridiagonal matrix block form.

$$Q = \begin{pmatrix} A_1^{(0)} & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \\ & & & & \end{pmatrix}$$

with matrices $A_0, A_1^{(0)}, A_1, A_2 \in \mathbb{R}^{K \times K}$. The generator Q is assumed to be irreducible. We call this system a level-independent QBD-process with finite number of phases.

We define the generator

$$A := A_2 + A_1 + A_0$$

which plays an important role in stability analysis of QBD processes.

Definition 2.1.25. For many of the statements in this Section 2.1.3 it is convenient to partition $\pi \in [0,1]^{\mathbb{N}_0 \times K}$ – the stochastic solution of the equation $\pi Q = 0$ – by levels

$$\pi = (\pi^{(0)}, \pi^{(1)}, \pi^{(2)}, \dots) \tag{2.1.41}$$

with

$$\pi^{(n)} := (\pi(n,k) : k \in K), \quad n \in \mathbb{N}_0. \tag{2.1.42}$$

Definition 2.1.26 (Loss system as level-independent QBD with finite number of phases)**.** Consider a loss system $Z = (X(t), Y(t) : t \in \mathbb{R}_0^+)$ as defined in Section 2.1.1 with irreducible generator Q, finite set K of environment states, constant arrival rate λ, and constant service rate μ. Then we have a special case of QBD defined in Definition 2.1.24 with

$$A_1^{(0)} = -\lambda I_W + V,$$

$$A_2 = \mu I_W R, \qquad A_1 = -(\mu+\lambda) I_W + V, \qquad A_0 = \lambda I_W$$

$$\text{and } A := \mu I_W R - (\mu+\lambda) I_W + V + \lambda I_W = \mu I_W (R - I) + V.$$

Lemma 2.1.27. *Given a loss system as in Definition 2.1.26, then the matrix A is irreducible.*

Proof. For any $m \in K$ and any $k \in K$ with $m \neq k$ it holds

$$(I_W(R-I)+V)_{km} > 0 \Longleftrightarrow \underbrace{\left(\mu I_W(R-I)+V\right)_{km}}_{=A_{km}} > 0.$$

Therefore matrix A is irreducible if and only if matrix $I_W(R-I)+V$ is irreducible. The irreducibility of $I_W(R-I)+V$ follows from irreducibility of Q according to Proposition 2.1.9. □

Proposition 2.1.28. *A loss system as defined in Definition 2.1.26 is positive recurrent if and only if $\mu > \lambda$, it is null recurrent if and only if $\mu = \lambda$, and it is transient if and only if $\mu < \lambda$.*

Proof. According to Lemma 2.1.27 the matrix A is irreducible, therefore we can apply [LR99, Theorem 7.2.4], which says, that the process is positive recurrent if and only if

$$xA_2\mathbf{e} > xA_0\mathbf{e}$$

where x is the unique stochastic solution of the equation $xA = 0$. For the proposition statement we do not need to know x exactly, it is sufficient to know that it is stochastic and positive. It is positive because A is irreducible and finite. We substitute $A_2 = \mu I_W R$ and $A_0 = \lambda I_W$ in the inequality above

$$x\mu I_W \underbrace{R\mathbf{e}}_{=\mathbf{e}} > x\lambda I_W\mathbf{e} \Longleftrightarrow \mu\,\underbrace{(xI_W\mathbf{e})}_{>0} > \lambda\,\underbrace{(xI_W\mathbf{e})}_{>0}.$$

Similarly according to [LR99, Theorem 7.2.4] the process is null recurrent if and only if

$$xA_2\mathbf{e} = xA_0\mathbf{e} \Longleftrightarrow \mu = \lambda$$

and it is transient if and only if

$$xA_2\mathbf{e} < xA_0\mathbf{e} \Longleftrightarrow \mu < \lambda.$$

□

Loss systems defined in Definition 2.1.26 are special case of level-dependent QBD with product form steady-state probability, where level is independent from phase, analyzed in [LR99]. We establish a link between our results and a more general result [LR99, Theorem 15.1.3]. The following Theorem 2.1.29 is from [LR99, Theorem 15.1.3] adopted to our notations and definitions.

Theorem 2.1.29 ([LR99, Theorem 15.1.3]). *The continuous time QBD is positive recurrent, and the level and phase are independent in steady state if and only if there exists a positive vector θ with $\theta\mathbf{e} = 1$ and a positive scalar η with $\eta < 1$ such that*

$$\theta(A_0 + \eta A_1 + \eta^2 A_2) = 0 \tag{2.1.43}$$

and

$$\theta(A_1^{(0)} + \eta A_2) = 0. \tag{2.1.44}$$

Then, we have that $\pi^{(n)} = (1 - \eta)\eta^n\theta$.

Proof. See [LR99, proof of Theorem 15.1.3]. □

For the loss system from Definition 2.1.24 we have $\eta = \frac{\lambda}{\mu}$. The value η is less than 1 if and only if $\lambda < \mu$ which is the criterion for positive recurrence in Proposition 2.1.28, as well as the summability condition (2.1.5) in Theorem 2.1.5.

The matrix $(A_1^{(0)} + \eta A_2)$ in (2.1.44) is matrix Q_{red} in (2.1.18):

$$A_1^{(0)} + \eta A_2 = -\lambda I_W + V + \frac{\lambda}{\mu}\mu I_W R = \lambda I_W(R - I) + V = Q_{\text{red}}.$$

The matrix $A_0 + \eta A_1 + \eta^2 A_2$ is $\frac{\lambda}{\mu}Q_{\text{red}}$ with $\eta = \frac{\lambda}{\mu}$:

$$\begin{aligned} A_0 + \eta A_1 + \eta^2 A_2 =& \lambda I_W + \frac{\lambda}{\mu}(-(\mu + \lambda)I_W + V) + \frac{\lambda^2}{\mu^2}\mu I_W R \\ =& \frac{\lambda}{\mu}(\lambda I_W(R - I) + V) = \frac{\lambda}{\mu}Q_{\text{red}}. \end{aligned}$$

Therefore θ in Theorem 2.1.29 and θ in Theorem 2.1.5 are the same.

Now we analyze the general case, where λ and μ may depend on the number n of customers and K can be infinite.

Definition 2.1.30. Let $Z = (X(t), Y(t) : t \in \mathbb{R}_0^+)$ be a level-dependent quasi-birth-death process on a state space $\mathbb{N}_0 \times K$ where $X(t) \in \mathbb{N}_0$ describes a level and $Y(t) \in K$ describes a phase. K is a countable set. Let Q be the generator of Z which has tridiagonal matrix block form

$$Q = \begin{pmatrix} A_1^{(0)} & A_0^{(0)} & & & \\ A_2^{(1)} & A_1^{(1)} & A_0^{(1)} & & \\ & A_2^{(2)} & A_1^{(2)} & A_0^{(2)} & \\ & & \ddots & \ddots & \ddots \\ & & & & \end{pmatrix}$$

with matrices $A_0^{(n)}, A_1^{(n)}, A_2^{(n)} \in \mathbb{R}^{K \times K}$ for all n.

Definition 2.1.31. Consider a loss system $Z = (X(t), Y(t) : t \in \mathbb{R}_0)$ as defined in Section 2.1.1 with generator Q, countable set of environment states K, arrival rates $\lambda(n)$, and service rates $\mu(n)$ where n is the current number of customers in the system or the current level in terms of QBD-processes. Then we have a special case of QBD from Definition 2.1.30 with

$$\begin{aligned} &A_1^{(0)} := -\lambda(0)I_W + V, \quad A_0^{(0)} = \lambda(0)I_W, \\ &A_2^{(n)} := \mu(n)I_W R, \qquad A_1^{(n)} = -(\mu(n)+\lambda(n))I_W + V, \quad A_0^{(n)} = \lambda(n)I_W, \qquad n \geq 1. \end{aligned} \tag{2.1.45}$$

The following Proposition 2.1.35 is generalization of [RT96, Corollary 2.1 and Corollary 2.2] *without* irreducibility Assumption 2.1.32. But before going to Proposition 2.1.35 we will summarize some important results from [RT96] adapting them to our notations and definitions.

Assumption 2.1.32 (Irreudiciblity assumption from [RT96, p. 123])**.**

$$A_1^{(0)} + A_0^{(0)} \text{ is irreducible} \tag{2.1.46}$$

and

$$A_0^{(n)} + A_1^{(n)} + A_2^{(n)} \text{ are irreducible for all } n \geq 0. \tag{2.1.47}$$

Proposition 2.1.33 ([RT96, pp. 123–124])**.** *Given a regular, irreducible continuous time Markov process $Z(t)$ as in Definition 2.1.30 and assume that the irreducibility Assumption 2.1.32 is satisfied. Then $Z(t)$ is positive recurrent if and only if the system of equations*

$$x_0 \left[A_1^{(0)} + \mathcal{R}_0 A_2^{(1)}\right] = 0 \tag{2.1.48}$$

has a positive solution x_0 such that

$$x_0 \left[\sum_{n=0}^{\infty} \prod_{\ell=0}^{n-1} \mathcal{R}_\ell\right] \mathbf{e} < \infty \tag{2.1.49}$$

where the family of matrices $(\mathcal{R}_n)_{n \in \mathbb{N}_0}$ is the minimal non-negative family that satisfies the equations

$$A_0^{(n)} + \mathcal{R}_n A_1^{(n+1)} + \mathcal{R}_n \mathcal{R}_{n+1} A_2^{(n+2)} = 0, \qquad \forall n \in \mathbb{N}_0. \tag{2.1.50}$$

Then, when conditions 2.1.48 and 2.1.49 are satisfied, the stationary distribution π is of the form

$$\pi^{(n)} = \pi^{(0)} \prod_{\ell=0}^{n-1} \mathcal{R}_\ell \tag{2.1.51}$$

with the vector $\pi^{(0)}$ the same as x_0 but normalized so that $\pi \mathbf{e} = 1$.

Proof. For the proof, in [RT96, p. 123] the authors refer to [BT95]. □

Remark 2.1.34. In the paper [BT95] preceding [RT96], the authors do not require Assumption 2.1.32. Also, in [RT96], the Assumption 2.1.32 is mentioned only once but is not referred to explicitly later. The purpose of Assumption 2.1.32 in [RT96] is not clear. If it would be possible to remove Assumption 2.1.32 from Proposition 2.1.33, this new proposition would cover most of the product form systems in this thesis. On the other side many of our main examples violate the Assumption 2.1.32.

The original proofs of [RT96, Corollary 2.1 and Corollary 2.2] rely on Proposition 2.1.33 and use direct or indirect the matrix family $(\mathcal{R}_n)_{n\in\mathbb{N}_0}$. In contrast to [RT96] we will use neither results from Proposition 2.1.33 nor the matrix family $(\mathcal{R}_n)_{n\in\mathbb{N}_0}$, but instead, we will directly substitute a product form for π into equation $\pi Q = 0$.

Proposition 2.1.35. *Given a regular irreducible quasi-birth-death process Z as in Definition 2.1.30. Then the following statements are equivalent:*

(a) Z is positive recurrent and its steady-state distribution π has a product form $\pi(n,k) = \xi(n)\theta(k)$ with positive stochastic vectors ξ and θ.

(b) There exists a positive stochastic solution θ of

$$\theta(A_1^{(0)} + \eta_0 A_2^{(1)}) = 0 \tag{2.1.52}$$

and positive numbers η_n, $n \in \mathbb{N}_0$ such that

$$\sum_{n=0}^{\infty}\prod_{\ell=0}^{n-1}\eta_\ell < \infty \tag{2.1.53}$$

and

$$\theta\left(A_0^{(n)} + \eta_n A_1^{(n+1)} + \eta_n\eta_{n+1}A_2^{(n+2)}\right) = 0 \qquad \forall n \geq 0. \tag{2.1.54}$$

If (b) holds, then the steady-state distribution has the form

$$\pi^{(n)} = C^{-1}\left(\prod_{\ell=0}^{n-1}\eta_\ell\right)\theta \tag{2.1.55}$$

with normalization constant C^{-1}, i.e $\pi^{(n)} = \xi(n)\theta$ with $\xi(n) = C^{-1}\left(\prod_{\ell=0}^{n-1}\eta_\ell\right)$.

Proof. (a)⇒(b) Let $\pi(n,k) = \xi(n)\theta(k)$ with positive stochastic vectors ξ and θ be the product form solution of a positive recurrent system with generator Q as defined in Definition 2.1.30. Then π is positive and the equation $\pi Q = 0$ holds. Given with $\pi^{(n)} := \xi(n)\theta$ is for $n = 0$

$$\pi^{(0)}A_1^{(0)} + \pi^{(1)}A_2^{(1)} = 0 \Longleftrightarrow \theta A_1^{(0)} + \frac{\xi(1)}{\xi(0)}\theta A_2^{(1)} = 0. \tag{2.1.56}$$

The equation (2.1.56) is equivalent to (2.1.52) with $\eta_0 := \frac{\xi(1)}{\xi(0)}$ if we can use the distributive law, which is not always possible with any arbitrary infinite matrices $A_1^{(n)}$ and $A_2^{(n)}$ from $\mathbb{R}^{K\times K}$. But in our case we can do it, because every k-th element of the left side of the equation (2.1.56) is an absolute convergent series $\sum_{\ell\in K}\theta(\ell)A_{1(\ell,k)}^{(0)} + \sum_{\ell\in K}\theta(\ell)\frac{\xi(1)}{\xi(0)}A_{2(\ell,k)}^{(1)}$.

To prove the absolute convergence, we use the fact that the only negative element in the sum is $A^{(0)}_{1(k,k)}$

$$\begin{aligned}&\sum_{\ell\in K}\left|\theta(\ell)A^{(0)}_{1(\ell,k)}\right| + \sum_{\ell\in K}\left|\theta(\ell)\frac{\xi(1)}{\xi(0)}A^{(1)}_{2(\ell,k)}\right|\\ &= \underbrace{\left(\sum_{\ell\in K}\theta(\ell)A^{(0)}_{1(\ell,k)} + \sum_{\ell\in K}\theta(\ell)\frac{\xi(1)}{\xi(0)}A^{(1)}_{2(\ell,k)}\right)}_{=0} - 2\theta(k)\frac{\xi(1)}{\xi(0)}A^{(1)}_{1(k,k)} < \infty.\end{aligned}$$

For all $n \geq 0$ we have

$$\begin{aligned}&\pi^{(n)}A_0^{(n)} + \pi^{(n+1)}A_1^{(n+1)} + \pi^{(n+2)}A_2^{(n+2)} = 0\\ \iff\ & \theta A_0^{(n)} + \theta\frac{\xi(n+1)}{\xi(n)}A_1^{(n+1)} + \theta\frac{\xi(n+2)}{\xi(n)}A_2^{(n+2)} = 0 \qquad (2.1.57)\\ \iff\ & \theta\left(A_0^{(n)} + \frac{\xi(n+1)}{\xi(n)}A_1^{(n+1)} + \frac{\xi(n+2)}{\xi(n)}A_2^{(n+2)}\right) = 0. \qquad (2.1.58)\end{aligned}$$

For the equivalence between (2.1.57) and (2.1.58) we use distributive law with the similar argumentation as for $n = 0$: Every k-th element of the right side of the equation (2.1.57) is an absolute convergent series $\sum_{\ell\in K}\theta(\ell)A^{(n)}_{0(\ell,k)} + \sum_{\ell\in K}\theta(\ell)\frac{\xi(n+1)}{\xi(n)}A^{(n+1)}_{1(\ell,k)} + \sum_{\ell\in K}\theta(\ell)\frac{\xi(n+2)}{\xi(n)}A^{(n+2)}_{2(\ell,k)}$ because the only negative element in the sum is $A^{(n+1)}_{1(k,k)}$. We estimate

$$\begin{aligned}&\sum_{\ell\in K}\left|\theta_\ell A^{(n)}_{0(\ell,k)}\right| + \sum_{\ell\in K}\left|\theta_\ell\frac{\xi(n+1)}{\xi(n)}A^{(n+1)}_{1(\ell,k)}\right| + \sum_{\ell\in K}\left|\theta_\ell\frac{\xi(n+2)}{\xi(n)}A^{(n+2)}_{2(\ell,k)}\right|\\ &= \underbrace{\left(\sum_{\ell\in K}\theta_\ell A^{(n)}_{0(\ell,k)} + \sum_{\ell\in K}\theta_\ell\frac{\xi(n+1)}{\xi(n)}A^{(n+1)}_{1(\ell,k)} + \sum_{\ell\in K}\theta_\ell\frac{\xi(n+2)}{\xi(n)}A^{(n+2)}_{2(\ell,k)}\right)}_{=0}\\ &\quad -2\theta_k\frac{\xi(n+1)}{\xi(n)}A^{(n+1)}_{1(k,k)} < 0.\end{aligned}$$

Setting $\eta_n := \frac{\xi(n+1)}{\xi(n)}$ in (2.1.58) we get the equations (2.1.54).

Due to construction of η_n from positive stochastic vector ξ, the summability condition (2.1.53) holds

$$\sum_{n=0}^{\infty}\prod_{\ell=0}^{n-1}\eta_\ell = \sum_{n=0}^{\infty}\prod_{\ell=0}^{n-1}\frac{\xi(\ell+1)}{\xi(\ell)} = \frac{1}{\xi(0)}\underbrace{\sum_{n=0}^{\infty}\xi(\ell)}_{=1} < 0.$$

Finally we show that vector $\pi^{(n)}$ can be represented in the form (2.1.55) with proposed η_n:

$$\pi^{(n)} = \xi(n)\theta = \underbrace{\xi(0)}_{=:C^{-1}}\prod_{\ell=0}^{n-1}\eta_\ell\theta.$$

(b)⇒(a) Assume we have a regular irreducible process with generator Q as defined Definition 2.1.30 such that there exist positive η_n and positive stochastic vector θ such that the equations (2.1.52) and (2.1.54) hold and the summability condition (2.1.53) is satisfied. Then $\pi := \left(\pi^{(n)} : n \in \mathbb{N}_0\right)$ with $\pi^{(n)} := C^{-1}\left(\prod_{\ell=0}^{n-1}\eta_\ell\right)\theta$ for all n with normalization constant $C = \sum_{n=0}^{\infty}\left(\prod_{\ell=0}^{n-1}\eta_\ell\right)$ is a positive probability measure and it solves the equations

$$\pi^{(0)}A_1^{(0)} + \pi^{(1)}A_2^{(1)} = 0 \Longleftrightarrow C^{-1}\theta A_1^{(0)} + C^{-1}\eta_0\theta A_2^{(1)} = 0$$

and

$$\pi^{(n)}A_0^{(n)} + \pi^{(n+1)}A_1^{(n+1)} + \pi^{(n+2)}A_2^{(n+2)} = 0$$
$$\Longleftrightarrow \left(C^{-1}\prod_{\ell=0}^{n-1}\eta_\ell\right)\theta A_0^{(n)} + \left(C^{-1}\prod_{\ell=0}^{n-1}\eta_\ell\right)\eta_n\theta A_1^{(n+1)} + \left(C^{-1}\prod_{\ell=0}^{n-1}\eta_\ell\right)\eta_n\eta_{n+1}\theta A_2^{(n+2)} = 0.$$

Because Q is irreducible the proposed positive stochastic solution π is the unique stationary probability of the process. Because π is finite and positive, the process is positive recurrent. □

Remark 2.1.36. In [RT96], all the subsequent corollaries and theorems are derived from or based on [RT96, Theorem 2.1] and their proofs can be traced back to the equations and inequality in Proposition 2.1.33 where these equations are said to be valid under irreducibility Assumption 2.1.32. In loss systems, irreducibility of $A_1^{(0)} + A_0^{(0)}$ is equivalent to irreducibility of matrix V:

$$A_1^{(0)} + A_0^0 = (V - \lambda(0)I_W) + \lambda(0)I_W = V.$$

Often the matrix V is not irreducible, see for example inventory models with (r, Q) policy in Example 2.2.2 and (r, S) policy in Example 2.2.3, systems with finite buffer in Section 2.2.3, and unreliable systems with $N \geq 1$ in Section 2.2.4.

Example 2.1.37. A special case of a loss system with reducible V is a tandem system of two $M/M/1/\infty$ nodes represented as a loss system where the first node is modeled by the queue and the second one, with rate ν, is modeled as an environment. The corresponding generator V has the form

$$V = \left(\begin{array}{c|ccc} & 0 & 1 & \dots \\ \hline 0 & 0 & 0 & \\ 1 & \nu & -\nu & \\ \vdots & & \ddots & \ddots \end{array}\right).$$

Remark 2.1.38. An ergodic loss system is a special case of systems with product-form steady-state distributions described in Proposition 2.1.35. When we substitute $A_1^{(0)}$ and $\eta_0 A_2^{(1)}$ from (2.1.45) and $\eta_0 := \frac{\lambda(0)}{\mu(1)}$ into (2.1.52) we get

$$\theta\left(-\lambda(0)I_W + V + \frac{\lambda(0)}{\mu(1)}\mu(1)I_W R\right) = 0$$

which corresponds to the equation $\theta Q_{\text{red}}(0) = 0$ in Theorem 2.1.5. Substituting $A_2^{(n)}$, $A_1^{(n)}$, $A_0^{(n)}$ from (2.1.45) and $\eta_n := \frac{\lambda(n)}{\mu(n+1)}$ into (2.1.54) leads to

$$\begin{aligned}
\theta\Bigg(\lambda(n)I_W + \frac{\lambda(n)}{\mu(n+1)}\Big(- \big(\mu(n+1)+\lambda(n+1)\big)I_W + V\Big)& \\
+\frac{\lambda(n)}{\mu(n+1)}\frac{\lambda(n+1)}{\mu(n+2)}\mu(n+2)I_W R\Bigg) &= 0, \qquad \forall n \geq 0 \\
\iff \frac{\lambda(n)}{\mu(n+1)}\big(\underbrace{\lambda(n+1)I_W(R-I)+V}_{=Q_{\text{red}}(n+1)}\big) &= 0, \qquad \forall n \geq 0
\end{aligned}$$

which is equivalent to the equations $\theta Q_{\text{red}}(n) = 0$ for all $n \geq 1$ in Theorem 2.1.5.

The condition (2.1.53) is the summability condition (2.1.5) in Theorem 2.1.5.

Remark 2.1.39 (Special case with constant λ, $K_W = K$ and irreducible V). A special case of loss systems in Definition 2.1.31 with constant λ, $K_W = K$ and irreducible V is a special case of the system in [RT96, Corollary 2.3]. Constant λ is required to keep the matrices $A_0^{(n)}$ constant for all n and $K_W = K$ is necessary to keep the matrices $A_2^{(n)}$ stochastic. Irreducibility of V is required by Remark 2.1.36.

2.1.4. Approximation of systems with no loss

We consider an $M/M/1/\infty$ queue in a non-autonomous random environment with service interruptions due to environment conditions. The queue may change the environment state each time a served customer leaves the system. In case of service interruption there is NO customer loss on arrival and newly arriving customers join the queue. A simple version of such a system is an $M/M/1/\infty$ queue with breakdowns, which has an autonomous environment with only two states – blocking and non-blocking. See [CM96, p. 269] for application of this principle for Jackson networks.

In [WC58, pp. 90–91], White and Christie derived the steady-state distribution of queue length of a $M/M/1/\infty$ system with exponential breakdown and repair times. In [AN63, Model A], Avi-Itzhak and Naor analyzed a system with general service and repair times and derived formulas for average waiting time and average queue length. In [MA68], Mitrani and Avi-Itzhak analyzed $M/M/N/\infty$ systems with exponential repair times and gave explicit formulas for the moment generating function of the queue size for $N \leq 2$. According to Chakka and Mitrani in [CM96, p. 269], the finding of the joint or marginal distribution of a queue size was still an open problem for $N > 1$ at that time.

A classical way to approximate a queue with breakdowns in an autonomous environment is the *reduced work-rate approximation*. For this method one first calculates breakdown probability for the environment, then replaces the original queue with environment by a queue with a smaller service rate without environment. The smaller service rate is the original service rate multiplied by the complement of the breakdown probability. It is based on an idea that the overall service capacity of both systems remains the same. See for example [CM96, pp. 267–269].

The aim of this section is to create an approximation methods which uses loss systems. In contrast to the reduced work-rate approximation in our approximation method we will keep the service rate, but adjust the throughput instead. We will also allow the environment to be non-autonomous and the resulting approximation will contain both, the queue and the environment.

We will focus our analysis on systems with constant input rate λ, constant service rate μ, and finite environment state space. This class of systems contains systems like the previously mentioned system with an autonomous two-state environment, a tandem system with finite buffer in Section 2.2.6, and inventory systems with (r, Q)-policy and backordering [SD05].

Notation We will use subscript NL (**N**o **L**oss) for the systems, whose customers are not lost due to blocking, in order to distinguish them from loss systems, whose customers are lost due to blocking.

Definition 2.1.40. Consider a non-explosive irreducible continuous-time Markov process $Z_{\mathrm{NL}} := ((X_{\mathrm{NL}}(t), Y_{\mathrm{NL}}(t)) : t \geq 0)$ describing a queueing system $X_{\mathrm{NL}}(t) \in \mathbb{N}_0$ in a random environment $Y_{\mathrm{NL}}(t) \in K$ on a finite state space $K = K_W \uplus K_B$ with irreducible generator $Q_{\mathrm{NL}} = (q_{\mathrm{NL}}((n,k),(n',k')) : n, n' \in \mathbb{N}, k, k' \in K)$:

$$\begin{aligned} q_{\mathrm{NL}}((n,k),(n+1,k)) &= \lambda_{\mathrm{NL}}, && k \in K, \\ q_{\mathrm{NL}}((n,k),(n-1,m)) &= \mu R(k,m), && k \in K_W, n > 0, \\ q_{\mathrm{NL}}((n,k),(n,m)) &= v(k,m) \in \mathbb{R}_0^+, && k \neq m, \\ q_{\mathrm{NL}}((n,k),(i,m)) &= 0, && \text{otherwise for } (n,k) \neq (i,m). \end{aligned}$$

with constant arrival rate $\lambda_{\mathrm{NL}} \in \mathbb{R}^+$, constant service rate $\mu \in \mathbb{R}^+$, non-empty set $K_W \subset K$, stochastic matrix R, and generator $V := (v(k,m) : k, m \in K)$. We call this kind of systems a system *with no loss* or system *without loss*.

Remark 2.1.41. Note that for the system defined in Definition 2.1.40 the irreducibility of the whole system is equivalent to irreducibility of the matrix $I_W(R - I) + V$.

The system in Definition 2.1.40 can be interpreted as

1. an $M/M/1/\infty$ queue X_{NL} in a random environment Y_{NL}, or as
2. a level-independent QBD process where X_{NL} describes the level and Y_{NL} describes the phase.

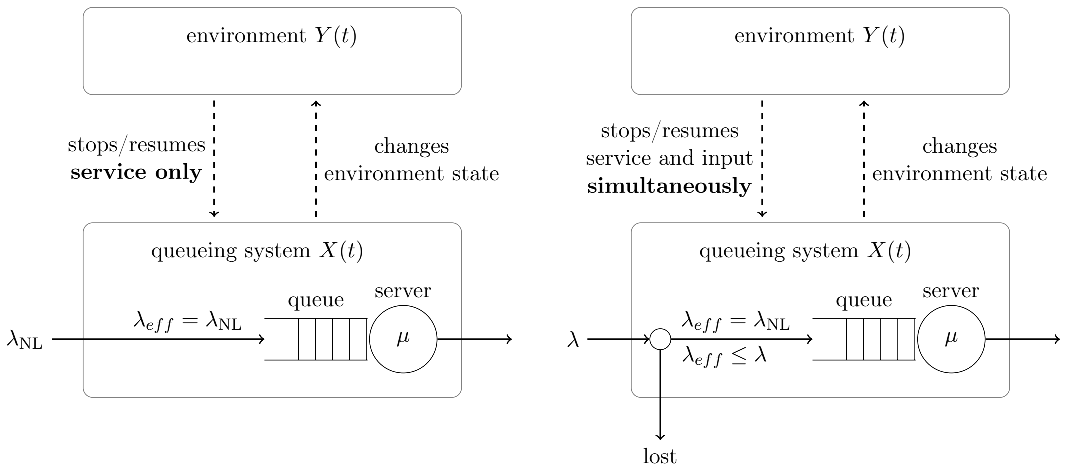
We will use the first interpretation to construct the associated loss system and the second interpretation for stability analysis.

2.1.4.1. Stability of the model with no loss

We introduce a function f which connects the steady-state analysis of system with no loss and system with loss having the same parameters K, K_B, μ, V and R.

Definition 2.1.42. For given finite state space K, stochastic matrix $R \in [0,1]^{K \times K}$, generator $V \in \mathbb{R}^{K \times K}$, and non-empty set of working states $K_W \subset K$, such that $I_W(R - I) + V$ is irreducible, we define a function

$$\begin{aligned} f : \mathbb{R}^+ &\longrightarrow (0,1]^K \\ x &\mapsto f(x) \end{aligned}$$



(a) System with no loss.

(b) Loss-system approximation.

Figure 2.1.1.: A system with no loss approximated by a loss system

which maps a positive value x to the unique stochastic solution $f(x)$ of the equation

$$f(x)\,(xI_W(R-I)+V) = 0. \tag{2.1.59}$$

The function f is well defined:

- For any $x \in \mathbb{R}^+$, any entry of matrix $xI_W(R-I)+V$ is positive (negative, zero) if and only if an entry with the same position in matrix $I_W(R-I)+V$ is positive (negative, zero). Further $I_W(R-I)+V$ is irreducible by definition. Therefore $xI_W(R-I)+V$ is irreducible.
- For any $x \in \mathbb{R}^+$, matrix $xI_W(R-I)+V$ is a generator, because of $(xI_W(R-I)+V)\mathbf{e} = xI_W \underbrace{(R\mathbf{e} - I\mathbf{e})}_{=0} + \underbrace{V\mathbf{e}}_{=0} = 0.$

Because $xI_W(R-I)+V$ is irreducible for any x, there exists a unique strictly positive stochastic solution $f(x)$ of (2.1.59).

Remark 2.1.43. The vector $f(\mu)$ plays an important role in stability analysis of systems without loss, as we will show in the following Proposition 2.1.45. It also holds $f(\lambda) = \theta$ where θ is the marginal distribution of the environment of the corresponding loss system with arrival rate λ, the same matrices R and V, and the same subset K_B as the system without loss. That means, we may reuse many of our results for loss systems – for example inventory models in Section 2.2.1 or systems with finite buffer in Section 2.2.3 – for stability analysis of corresponding systems with no loss. To do this, we only need to replace λ in θ by x and we will get $f(x)$ for free.

Lemma 2.1.44. *Function f from Definition 2.1.42 is continuous.*

Proof. Let x be an arbitrary fixed strictly positive value $x \in \mathbb{R}^+$. We define

$$G := (g_{km})_{k,m\in K} := xI_W(R-I)+V$$

and

$$\tilde{G} := (\tilde{g}_{km})_{k,m \in K} = (x+\delta) I_W (R - I) + V.$$

Both generators G and $\tilde{G}$ are irreducible.

We estimate the distance between any non-diagonal entry of the matrix $\tilde{G}$ and the corresponding entry of the matrix G for some small δ with $|\delta| < x$.

$$\begin{aligned}
|\tilde{g}_{km} - g_{km}| &= |1_{[k \in K_W]}(k) \delta r(k,m)| \\
&\leq \frac{|\delta|}{x} \Big| 1_{[k \in K_W]}(k) x r(k,m) + 1_{[k \in K_W]}(k) \underbrace{v(k,m)}_{\geq 0} \Big| \\
&= \frac{|\delta|}{x} g_{km} \qquad \forall m \neq k.
\end{aligned}$$

According to [O'C93, Corollary 1] it holds for stochastic solution $f(x)G = 0$ and $f(x+\delta)\tilde{G} = 0$

$$|f(x+\delta)(k) - f(x)(k)| \leq \left(\left(\frac{1 + \frac{|\delta|}{x}}{1 - \frac{|\delta|}{x}} \right)^{|K|} - 1 \right) f(x)(k) \qquad \forall k \in K.$$

The limit for $\delta \to 0$ of the right side of the inequality is zero for any $x \in \mathbb{R}^+$and any $k \in K$ therefore f is continuous. □

Proposition 2.1.45. *Given the Markov process $Z_{NL} := ((X_{NL}(t), Y_{NL}(t)) : t \geq 0)$ with no loss as defined in Definition 2.1.40. Then the system is positive recurrent if and only if*

$$\mu \sum_{k \in K_W} f(\mu)(k) > \lambda_{NL},$$

it is null recurrent if and only if

$$\mu \sum_{k \in K_W} f(\mu)(k) = \lambda_{NL},$$

and it is transient if and only if

$$\mu \sum_{k \in K_W} f(\mu)(k) < \lambda_{NL}.$$

Proof. The generator Q_{NL} has the form

$$Q_{\mathrm{NL}} = \begin{pmatrix} A_1^{(0)} & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \\ & & & & \end{pmatrix}$$

where

$$A_1^{(0)} = -\lambda_{\mathrm{NL}} I + V,$$

$$A_2 = \mu I_W R, \qquad A_1 = -\mu I_W - \lambda_{\mathrm{NL}} I + V, \qquad A_0 = \lambda_{\mathrm{NL}} I.$$

For $A := A_2 + A_1 + A_0$ it holds

$$A = \mu I_W R - \mu I_W - \lambda_{\mathrm{NL}} I + V + \lambda_{\mathrm{NL}} I = \mu I_W (R - I) + V. \tag{2.1.60}$$

From irreducibility of $I_W(R-I)+V$ follows irreducibility of A. According to [Neu81, Theorem 3.1.1] or [LR99, Theorem 7.2.3] the system is positive recurrent if and only if $\boldsymbol{\alpha} A_2 \mathbf{e} > \boldsymbol{\alpha} A_0 \mathbf{e}$ where $\boldsymbol{\alpha}$ is a stochastic solution of the equation $\boldsymbol{\alpha} A = 0$. With (2.1.60) and $f(\mu)$ defined as the stochastic solution of the equation $f(\mu)\left(\mu I_W(R-I)+V\right) = 0$ we immediately have $\boldsymbol{\alpha} = f(\mu)$. Therefore the system is positive recurrent if and only if

$$f(\mu) A_2 \mathbf{e} > f(\mu) A_0 \mathbf{e} \Longleftrightarrow f(\mu)\mu I_W \underbrace{R\mathbf{e}}_{=\mathbf{e}} > f(\mu)\lambda_{\mathrm{NL}} \underbrace{I\mathbf{e}}_{=\mathbf{e}} \Longleftrightarrow \mu \sum_{k \in K_W} f(\mu)(k) > \lambda_{\mathrm{NL}}.$$

Similarly, according to [LR99, Theorem 7.2.3], the system without loss is null recurrent if and only if

$$\boldsymbol{\alpha} A_2 \mathbf{e} = \boldsymbol{\alpha} A_0 \mathbf{e} \Longleftrightarrow \mu \sum_{k \in K_W} f(\mu)(k) = \lambda_{\mathrm{NL}}$$

and it is transient if and only if

$$\boldsymbol{\alpha} A_2 \mathbf{e} < \boldsymbol{\alpha} A_0 \mathbf{e} \Longleftrightarrow \mu \sum_{k \in K_W} f(\mu)(k) < \lambda_{\mathrm{NL}}.$$

□

2.1.4.2. Loss system approximation

In this subsection we will investigate whether it is possible to construct for an ergodic no loss system from Definition 2.1.40 an ergodic loss system on the same state space $\mathbb{N}_0 \times K$ and with the same parameters μ, R, V, K_W and the same effective arrival rate as the system without loss.

Proposition 2.1.46. *For an ergodic system defined in Definition 2.1.40 there exists a corresponding ergodic loss systems as defined in Section 2.1.1 with adjusted λ, with the same parameters K, K_W, μ, R and V and with the same throughput in steady state. This is equivalent to have the same effective arrival rate in steady state, both equal to λ_{NL}:*

$$\lambda \sum_{n \in \mathbb{N}_0} \sum_{k \in K_W} \pi(n,k) = \lambda_{NL} \tag{2.1.61}$$

where $\pi(n,k)$ is the steady-state distribution of the loss system.

Proof. The main idea of the proof is that equation (2.1.61) can be written as $\lambda \sum_{k \in K_W} f(\lambda)(k) = \lambda_{\mathrm{NL}}$. Using the intermediate value theorem we will show that there exists λ which solves the equation $\lambda \sum_{k \in K_W} f(\lambda)(k) = \lambda_{\mathrm{NL}}$ and that the loss system with an input rate λ is ergodic.

The system without loss is ergodic, therefore it is positive recurrent. According to Proposition 2.1.45 it holds

$$\mu \sum_{k \in K_W} f(\mu)(k) > \lambda_{\mathrm{NL}} > 0.$$

Together with

$$\lim_{x \to 0} x \underbrace{\sum_{k \in K_W} f(x)(k)}_{\leq 1} = 0$$

and continuity of the mapping $x \mapsto x \sum_{k \in K_W} f(x)(k)$ using intermediate value theorem we conclude the existence of some $\lambda \in (0, \mu)$ such that

$$\lambda \sum_{k \in K_W} \underbrace{f(\lambda)(k)}_{\theta(k)} = \lambda_{\mathrm{NL}}. \tag{2.1.62}$$

See Figure 2.1.2 on page 40.

We choose any λ satisfying (2.1.62) as arrival rate for the corresponding loss system. Because $\lambda < \mu$, the summability condition (2.1.5) from Theorem 2.1.5

$$C := \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu} \right)^n < \infty$$

is satisfied. The matrix $Q_{\mathrm{red}} = \lambda I_W (R - I) + V$ inherits its irreducibility property from A and has a unique positive stochastic solution $\theta = f(\lambda)$. According to Corollary 2.1.10 the corresponding loss system is ergodic with product form steady-state probability $\pi(n, k) = \xi(n)\theta(k)$. We multiply the left side of (2.1.62) with $\sum_n \xi(n) = 1$ and finally prove the equation (2.1.61):

$$\lambda \sum_n \xi(n) \sum_{k \in K_W} \theta(k) = \lambda_{\mathrm{NL}} \Longleftrightarrow \lambda \sum_n \sum_{k \in K_W} \underbrace{\xi(n)\theta(k)}_{=\pi(n,k)} = \lambda_{\mathrm{NL}}.$$

□

Proposition 2.1.47. *Given a system with no customer loss as in Definition 2.1.40 such that $x \sum_{k \in K_W} f(x)(k)$ is strictly monotone increasing in x. Then the following statements are equivalent:*

(a) The system without loss is ergodic.

(b) There exists a corresponding ergodic loss system with the same parameters K, K_W, μ, R and V and effective arrival rate in steady state λ_{NL}.

Furthermore the corresponding loss system is unique.

Remark 2.1.48. The strictly monotone increasing property of the function $x \sum_{k \in K_W} f(x)(k)$ can be interpreted as "the faster the service, the higher input rate a system without loss can process staying stable". This follows from condition $\mu \sum_{k \in K_W} f(\mu)(k) > \lambda$ for positive recurrence. Many systems have this monotonicity property – see for example Corollary 2.2.37 in Section 2.2.6 – but we can construct a system with no loss which penalize fast service by large delays as we will show in Example 2.1.49.

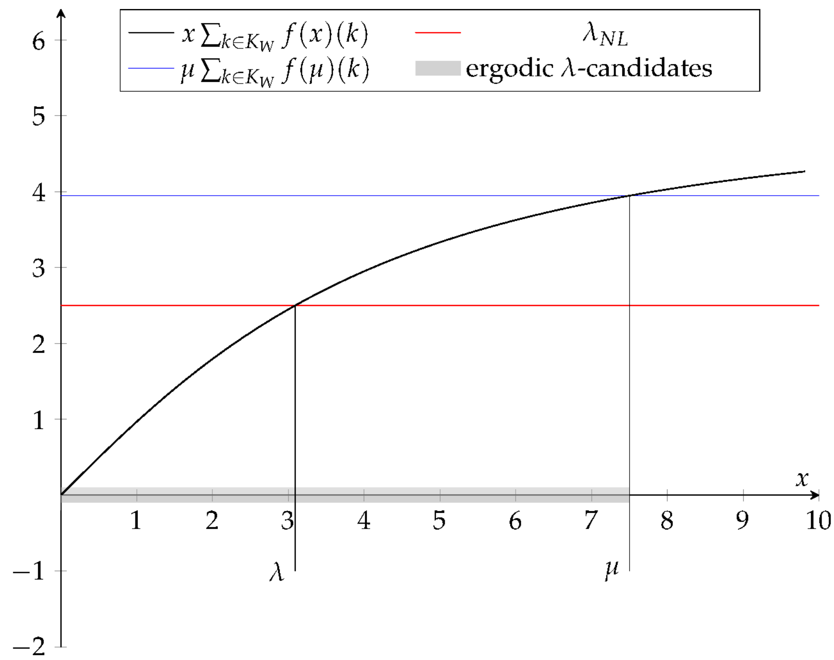


Figure 2.1.2.: Idea of the proof of Proposition 2.1.46. An example of $x \sum_{k \in K_W} f(x)(k)$ function for given parameters R, V and K_W. If parameters μ and λ_{NL} belong to an ergodic system with no loss, then the horizontal line $\mu \sum_{k \in K_W} f(\mu)(k)$ lies above the line λ_{NL}.

Proof of Proposition 2.1.47. (a)$\Longrightarrow$(b): According to Proposition 2.1.46, if the system without loss is ergodic then there exists a corresponding ergodic system with the same effective arrival rate. The uniqueness of the ergodic loss system follows from the requirement $\lambda \sum_{k \in K_W} \theta(k) = \lambda_{\mathrm{NL}}$ which is equivalent to $\lambda \sum_{k \in K_W} f(\lambda)(k) = \lambda_{\mathrm{NL}}$, see Remark 2.1.43. The last equation has a unique solution λ for given λ_{NL} due to strict monotonicity of $\lambda \sum_{k \in K_W} f(\lambda)(k)$.

(b)$\Longrightarrow$(a): We will prove this by contradiction.

Assume the corresponding system without loss is not ergodic but there exists a corresponding ergodic loss system with

$$\lambda \sum_{n \in \mathbb{N}_0} \sum_{k \in K_W} \pi(n,k) = \lambda \sum_{k \in K_W} \theta(k) = \lambda_{\mathrm{NL}} \tag{2.1.63}$$

and $\lambda < \mu$. According to Proposition 2.1.45 for all non-ergodic systems without loss it holds

$$\mu \sum_{k \in K_W} f(\mu)(k) \leq \lambda_{\mathrm{NL}} \overset{(2.1.63)}{\Longleftrightarrow} \mu \sum_{k \in K_W} f(\mu)(k) \leq \lambda \sum_{k \in K_W} \underbrace{\theta(k)}_{=f(\lambda)(k)} \overset{\text{str. mon. inc.}}{\Longleftrightarrow} \mu \leq \lambda.$$

So, we obtained a contradiction to $\mu > \lambda$. $\square$

The next example shows that strict monotonicity of $x \sum_{k \in K_W} f(x)(k)$ in x, required in Proposition 2.1.47, is not always given.

Example 2.1.49. Assume we have a system with three environment states: 1 - normal, 2 - sensible, and 3 - blocking.

After each service in normal state 1 the system immediately changes to the sensible state 2. After each service in the sensible state 2 the system immediately switches to the blocking state 3 and stops service.

The environment has an exponential distributed timer, which expires in state 2 with a fast rate α. After timeout the timer will put the environment from sensible state 2 back to the normal state 1. If the environment is in blocking state, it will change to a normal state with a very slow rate β.

We can model this system using following parameters: $K = \{1, 2, 3\}$, $K_W = \{1, 2\}$,

$$R = \left(\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & 1 & \\ 2 & & & 1 \\ 3 & & & 1 \end{array}\right), \qquad V = \left(\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & & \\ 2 & \alpha & -\alpha & \\ 3 & \beta & & -\beta \end{array}\right).$$

The irreducible matrix $xI_W(R - I) + V$ is

$$xI_W(R - I) + V = \left(\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & -x & x & \\ 2 & \alpha & -\alpha - x & x \\ 3 & \beta & & -\beta \end{array}\right).$$

For $f(x)$ the stochastic solution of the equation $f(x)\,(xI_W(R - I) + V) = 0$ it holds

$$f(x)(1) = \frac{\alpha + x}{x} \cdot \frac{\beta}{x} f(x)(3), \quad f(x)(2) = \frac{\beta}{x} f(x)(3), \quad f(x)(3) = \frac{x^2}{x^2 + 2\beta x + \alpha\beta}.$$

For $g(x) := x \sum_{k \in K_W} f(x)(k)$ it holds

$$g(x) = x\left(1 - \sum_{k \in K_B} f(x)(k)\right) = \frac{x\beta(2x + \alpha)}{x^2 + 2\beta x + \alpha\beta}.$$

Finally we choose the parameter $\alpha = 10$ (very fast) and $\beta = 0.1$ (very slow) in such a way, that the function $g(x)$ is not monotone. See Figure 2.1.3 on page 42.

2.1.4.3. Properties of the loss-system approximation

Consistency and small blocking probability If there are no blocking states in the system, i.e. $K_W = K$, a system without loss and input rate λ_{NL} defined in Definition 2.1.40 is a loss system. So, we would expect that our loss-system approximation with the same throughput λ_{NL} proposed in Proposition 2.1.46 is consistent and the approximation is exact. In fact, this is the case. By definition, the parameters K, K_W, μ, R and V remain the same, and for the adjusted λ it holds

$$\lambda \sum_{n \in \mathbb{N}_0} \sum_{k \in K_W} \pi(n, k) = \lambda_{\mathrm{NL}} \Longleftrightarrow \lambda \sum_{n \in \mathbb{N}_0, k \in K} \pi(n, k) = \lambda_{\mathrm{NL}} \Longleftrightarrow \lambda = \lambda_{\mathrm{NL}}.$$

That means, that the original system without loss and its loss-system approximation are identical.

From the practical point of view we can expect loss-system approximation to be good if the blocking probability of the original system without loss is small.

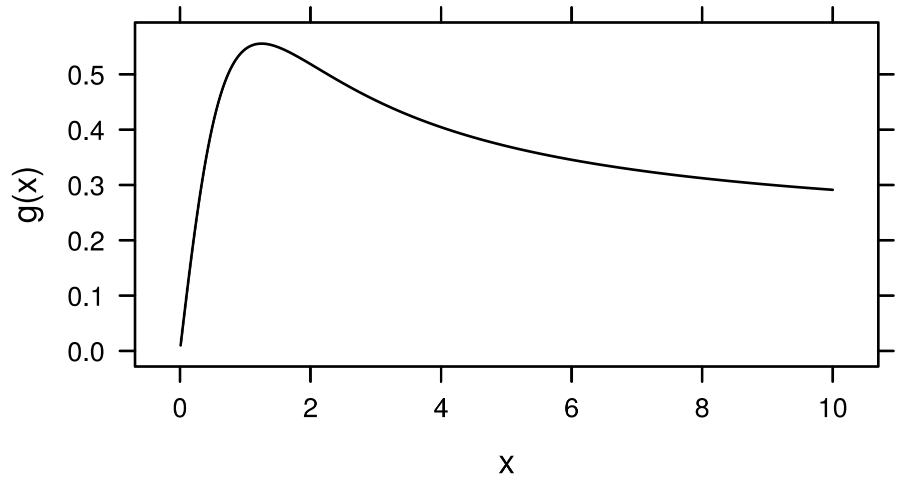


Figure 2.1.3.: $g(x) = x \sum_{k \in K_W} f(x)(k)$ from Remark 2.1.48 with with $\alpha = 10$ and $\beta = 0.1$.

Very fast service rate and non-negligible blocking probability An important example when the loss-system provides bad approximation is when someone tries to estimate the average number of customers in the system.

Definition 2.1.50. Given an value x and its approximation x_{appx}. Then we will call the value $x_{appx} - x$ the *signed absolute error*. And we will call the value $\frac{x_{appx}-x}{x}$ the *signed relative error*.

Assume we have a system with no loss, and non-empty set of blocking states, similar to a system defined in Definition 2.1.40, but with zero service time. In steady state, this system will have an empty queue when it is not blocked, but it can have a non-empty queue, when it is blocked. The average queue size in steady state $E(X_{\mathrm{NL},\mu=\infty})$ is strictly positive. If we will approximate this system with a loss system, then after a finite period of time, the loss system will have no customers. In this case the average queue size in steady state $E(X_{\mu=\infty})$ is zero.

The signed relative error$(E(X_{\mu=\infty}) - E(X_{\mathrm{NL},\mu=\infty}))\,/E(X_{\mathrm{NL},\mu=\infty})$ is then -1 and the absolute error is $-E(X_{\mathrm{NL},\mu=\infty})$.

From previous thoughts, we can expect that the loss-system approximation is not suitable to estimate the queue length when the service rate μ is large and the blocking probability is not negligible. Especially when the average size of the queue of the original system is expected to be significant larger than zero.

Insensitivity of θ Another important property of loss-system approximation is that the steady-state probability θ of the approximation does not depend on μ. See Remark 2.1.8. As a result, according to (2.1.62), the adjusted parameter λ

$$\lambda = \frac{\lambda_{NL}}{\sum_{k \in K_W} \theta(k)}$$

does not depend from μ too. We should be careful and take in account this property, if we want to approximate the environment steady-state probability of a system without loss, whose steady-state distribution of the environment significantly depends on μ.

2.2. Applications

2.2.1. Inventory models

In the following we describe an $M/M/1/\infty$-system with inventory management as it is investigated in [SSD$^+$06].

Definition 2.2.1. An $M/M/1/\infty$-system with inventory management is a single server with infinite waiting room under FCFS regime and an attached inventory.

There is a Poisson-λ-arrival stream, $\lambda > 0$. Customers request for an amount of service time which is exponentially distributed with mean 1. Service is provided with intensity $\mu > 0$.
The server needs for each customer exactly one item from the attached inventory. The on-hand inventory decreases by one at the moment of service completion. If the inventory is decreased to the reorder point $r \geq 0$ after the service of a customer is completed, a replenishment order is instantaneously triggered. The replenishment lead times are i.i.d. with distribution function $B = (B(t); t \geq 0)$. The size of the replenishment depends on the policy applied to the system. We consider two standard policies from inventory management, which lead to an M/M/1/∞-system with either (r, Q)-policy (size of the replenishment order is always $Q > r$) or with (r, S)-policy (replenishment fills the inventory up to maximal inventory size $S > r$).

During the time the inventory is depleted and the server waits for a replenishment order to arrive, no customers are admitted to join the queue ("lost sales").
All service, inter-arrival and lead times are assumed to be independent.
Let $X(t)$ denote the number of customers present at the server at time $t \geq 0$, either waiting or in service (queue length) and let $Y(t)$ denote the on-hand inventory at time $t \geq 0$. Then $\big((X(t), Y(t)), t \geq 0\big)$, the *queueing-inventory process* is a continuous-time Markov process for the $M/M/1/\infty$-system with inventory management. The state space of (X, Y) is $E = \{(n, k) : n \in \mathbb{N}_0, k \in K\}$, where $K = \mathbb{N}_0$ or $K = \{0, 1, \ldots, \kappa\}$, where $\kappa < \infty$ is the maximal size of the inventory at hand.

The system described above generalizes the lost sales case of classical inventory management where customer demand is not backordered but lost in case there is no inventory on hand (see Tersine [Ter94, p. 207]).

The general Theorem 2.1.5 produces as special application the following results on product-form steady-state distribution in integrated queueing inventory systems which are described in [SSD$^+$06].

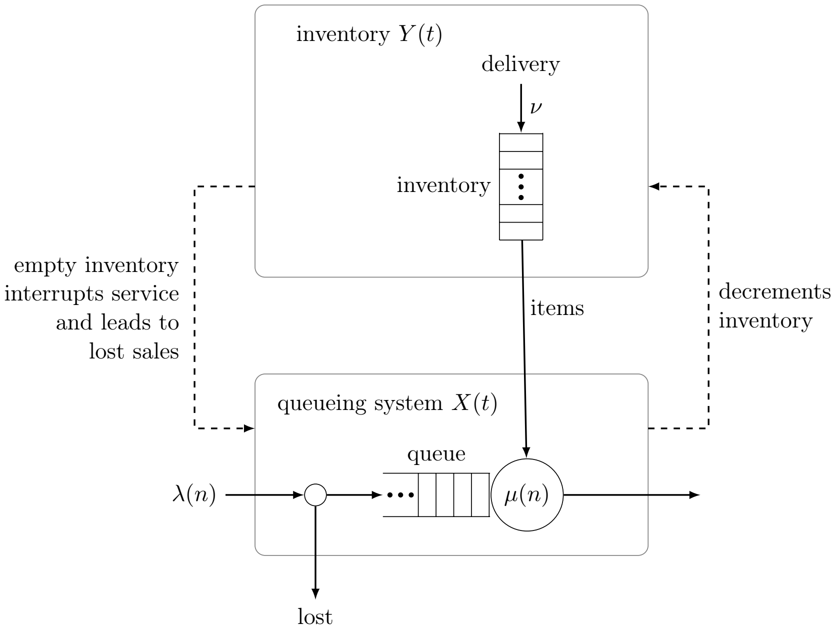
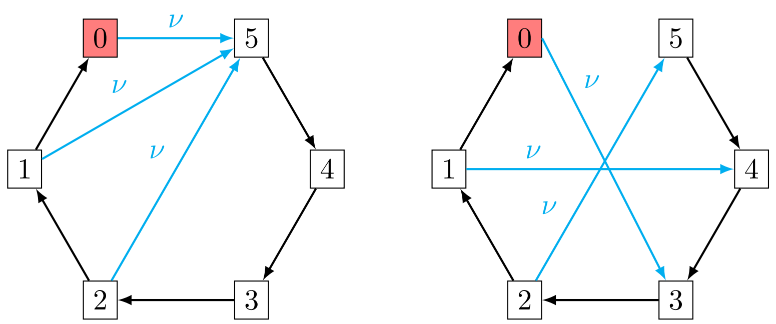


Figure 2.2.1.: $M/M/1/\infty$ inventory model with lost sales.



(a) $M/M/1/\infty$ system with $(r = 2, S = 5)$-policy and lost sales.

(b) $M/M/1/\infty$ system with $(r = 2, Q = 3)$-policy and lost sales.

Figure 2.2.2.: Environment transition and interaction diagram for lost sales environment systems. The environment process counts the number of items in inventory.

Example 2.2.2. [SSD+06] $M/M/1/\infty$ system with (r,S)-policy, exp(ν)-distributed lead times, and lost sales. The inventory management process under (r,S)-policy fits into the definition of the environment process by setting

$$K = \{0,1,\dots,S\}, \qquad K_B = \{0\},$$

$$R(0,0)=1, \quad R(k,k-1)=1, \quad 1 \le k \le S, \qquad v(k,m) = \begin{cases} \nu, & \text{if } 0 \le k \le r, m = S, \\ 0, & \text{otherwise for } k \ne m. \end{cases}$$

The queueing-inventory process (X,Y) is ergodic iff $\lambda < \mu$. The steady-state distribution $\pi = (\pi(n,k) : (n,k) \in E)$ of (X,Y) has product form

$$\pi(n,k) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\nu}\right)^n \theta(k),$$

where $\theta = (\theta(k) : k \in K)$ with normalization constant C is

$$\theta(k) = \begin{cases} C^{-1}(\frac{\lambda}{\nu}), & k = 0, \\ C^{-1}(\frac{\lambda+\nu}{\lambda})^{k-1}, & k = 1,\dots,r, \\ C^{-1}(\frac{\lambda+\nu}{\lambda})^{r}, & k = r+1,\dots,S. \end{cases} \tag{2.2.1}$$

Example 2.2.3. [SSD+06]$M/M/1/\infty$ system with (r,Q)-policy, exp(ν)-distributed lead times, and lost sales. The inventory management process under (r,Q)-policy fits into the definition of the environment process by setting

$$K = \{0,1,\dots,r+Q\}, \qquad K_B = \{0\},$$

$$R(0,0)=1, \quad R(k,k-1)=1, \quad 1 \le k \le S, \qquad v(k,m) = \begin{cases} \nu, & \text{if } 0 \le k \le r, m = k+Q, \\ 0, & \text{otherwise for } k \ne m. \end{cases}$$

The queueing-inventory process (X,Y) is ergodic iff $\lambda < \mu$. The steady-state distribution $\pi = (\pi(n,k) : (n,k) \in E$ of (X,Y) has product form

$$\pi(n,k) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\nu}\right)^n \theta(k),$$

where $\theta = (\theta(k) : k \in K)$ with normalization constant C is

$$\begin{aligned} \theta(0) &= C^{-1}\frac{\lambda}{\nu}, && \\ \theta(k) &= C^{-1}\left(\frac{\lambda+\nu}{\lambda}\right)^{k-1}, && k = 1,\dots,r, \\ \theta(k) &= C^{-1}\left(\frac{\lambda+\nu}{\lambda}\right)^{r}, && k = r+1,\dots,Q, \\ \theta(k+Q) &= C^{-1}\left(\frac{\lambda+\nu}{\lambda}\right)^{r} - \left(\frac{\lambda+\nu}{\lambda}\right)^{k-1}, && k = 1,\dots,r. \end{aligned}$$

Example 2.2.4. Recently Krishnamoorthy, Manikandan, and Lakshmy [KML13] analyzed an extension of the (r, S) and (r, Q) inventory systems with lost sales where the service time has a general distribution, and at the end of the service the customer receives with probability γ one item from the inventory while and with probability $(1 - \gamma)$ the inventory level stays unchanged. The authors calculate the steady-state distribution of the system, which has a product form, and give necessary and sufficient condition for stability. In the case of exponential service time our model from Section 2.1.1 encompasses this system.

Remark 2.2.5. At a first glance, it seems that the steady-state results from Example 2.2.2, Example 2.2.3 and Example 2.2.4 can be easily extended to models with multiple servers $M/M/m/\infty$ if we define $\mu(n) = \min(n, m)\mu$ where μ is a service rate of each server and if we assume $\sum_{n=0}^{\infty} C^{-1} \frac{\lambda^n}{\prod_{i=0}^{n-1} \mu(i+1)} < \infty$. According to Corollary 2.1.10, the extended versions of these inventory models with m servers, modeled as single server with stat dependent server rate, has the steady-state distribution

$$\pi(n, k) = C^{-1} \frac{\lambda^n}{\prod_{i=0}^{n-1} \mu(i+1)} \theta(k)$$

with the same θ as in the model with a single server. The problem we are faced with is weather this state dependent service model reflects our motivating production-inventory system. It turns out that this is not the case: Consider a system with two customers in the queue. When there are two items in the inventory, the total service rate is 2μ, but if there is only one item in the inventory, only one server can be active and the total service rate is μ. The service rate in this model depends on the environment in a more complex way than in the loss-system defined in Section 2.1.1. Therefore we cannot apply Corollary 2.1.10 or Theorem 2.1.5.

In [KMD15], the authors analyze a multiserver version of Example 2.2.4, they calculate a product-form steady-state distribution for a system with $m = 2$ servers but conjecture that there is no analytical solution for a system with $m \geq 3$ server.

Example 2.2.6. Recently Jung Woo Baek and Seung Ki Moon [BM14] analyzed an extension of the (r, Q) inventory system with lost sales, where they added an internal production process to the production-inventory system. Their system has two suppliers: the external one with (r, Q) policy and the internal one which always fills the inventory with a constant rate $\beta < \lambda$. The size of the inventory is unlimited. They calculated the steady-state distribution, which has a product form, and analyzed the long run costs. This model is a special case of a loss system introduced in Section 2.1.1.

Example 2.2.7. [6]This example is taken from [KN13], the notation is adapted to that used in Section 2.1.1: The authors study an inventory system under (r, S)-policy, which provides items for a server who processes and forwards the items in an on-demand production scheme. The processing time of each service is exponentially-μ distributed. The demand occurs in a Poisson-λ stream.

If demand arrives when the inventory is depleted it is rejected and lost to the system forever (lost sales).

[6]published as [KD15a, Example 7]

The complete system is a supply chain where new items are added to the inventory through a second production process which is interrupted whenever the inventory at hand reaches S. The production process is resumed each time the inventory level goes down to r and continues to be on until inventory level reaches S again. The times required to add one item into the inventory (processing time + lead time) when the production is on, are exponential-ν random variables.

All inter arrival times, service times, and production times are mutually independent.

For a Markovian description we need to record the queue length of not fulfilled demand ($\in \mathbb{N}_0$), the inventory on stock ($\in \{0, 1, \ldots, S\}$), and a binary variable which indicates when the inventory level is in $\{r+1, \ldots, S\}$ whether the second production process is on (=1) or off (= 0). (Note, that the second production process is always on, when the inventory level is in $\{0, 1, \ldots, r\}$, and is always off, when the inventory level is S.)

To fit this model into the framework of Section 2.1.1 we define a Markov process (X, Y) in continuous time with state space

$$E := \mathbb{N}_0 \times K, \text{ with } K := \{0, 1, \ldots, r\} \cup \{S\} \cup (\{r+1, \ldots, S-1\} \times \{0, 1\}) \text{ and } K_B = \{0\}.$$

The environment therefore records the inventory size and the status of the second production process, and blocking of the production system occurs due to stock out with lost sales regime.

Starting from Example 2.2.3, Saffari, Haji, and Hassanzadeh [SHH11] proved that under (r, Q) policy the integrated queueing-inventory $M/M/1/\infty$ system with hyper-exponential lead times (= mixtures of exponential distributions) has a product-form distribution. The proof is done by solving directly the steady-state equations. In [SAH13], Saffari, Asmussen, and Haji generalized this result to general lead time distributions. The proof of product form uses some intuitive arguments from related simplified systems and the marginal probabilities for the inventory position are derived using regenerative arguments.

In the following example we show that our models encompasses queueing-inventory systems with general replenishment lead times under (r, S) policy. This will allow us directly to conclude that for the ergodic system the steady state has product form and this will enable us to generalize the theorem (here Example 2.2.2) of [SSD$^+$06] to incorporate generally distributed lead times.

In a second step we will show that the results of Saffari, Haji, and Hassanzadeh [SHH11] and of Saffari, Asmussen, and Haji [SAH13] for queueing-inventory systems under (r, Q) policy can be obtained by our method as well and can even be slightly generalized.

We will consider lead time distributions of the following phase-type which are sufficient versatile to approximate any distribution on $\mathbb{R}_+$ arbitrary close.

Definition 2.2.8 (Phase-type distributions)**.** For $k \in \mathbb{N}$ and $\beta > 0$ let

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

denote the cumulative distribution function of the Γ–distribution with parameters β and k. Parameter k 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 $\Gamma_{\beta,k}$. ($\Gamma_{\beta,k}$ is called a k-stage Erlang distribution with shape parameter β.)

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

$$B(s) = \sum_{\ell=1}^{L} b(\ell)\Gamma_{\beta,\ell}(s), \quad s \geq 0, \tag{2.2.2}$$

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

To incorporate replenishment lead time distributions of phase-type we apply the supplemented variable technique. This leads to enlarging the phase space of the system, i.e. the state space of the inventory process Y. Whenever there is an ongoing lead time, i.e., when inventory at hand is less than $r+1$ we count the number of residual successive i.i.d. $\exp(\beta)$-distributed lead time phases which must expire until the replenishment arrives at the inventory.

The state space of (X, Y) then is $E = \mathbb{N}_0 \times K$ with

$$K = \{r+1, r+2, \dots, S\} \cup (\{0, 1, \dots, r\} \times \{L, \dots, 1\}),$$

and (X, Y) is irreducible on E.

Proposition 2.2.9. [7]*$M/M/1/\infty$ system with (r, S)-policy, phase-type replenishment lead time, state dependent service rates $\mu(n)$, and lost sales.*

The lead time distribution has a distribution function B from (2.2.2). *We assume that (X, Y) is positive recurrent and denote its steady-state distribution by*

$$\pi = (\pi(n, k) : n \in \mathbb{N}_0 \times K).$$

The steady-state distribution π of (X, Y) is of product form. With normalization constant C

$$\pi(n, k) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda}{\mu(i+1)} \cdot \theta(k) \tag{2.2.3}$$

where $\theta = (\theta(k) : k \in K)$ is for $r = 0$

$$\theta(j, \ell) = G^{-1} \left(\frac{\lambda + \beta}{\lambda}\right)^{j-1} \sum_{i=\ell}^{L} b(i) \left(\frac{\beta}{\lambda + \beta}\right)^{i-\ell} \binom{i - \ell + r - j}{r - j}, \quad j = 1, 2, \dots, r, \ell = 1, \dots, L, \tag{2.2.4}$$

$$\theta(0, \ell) = G^{-1} \frac{\lambda}{\beta} \left[\sum_{i=\ell}^{L} \left(\sum_{g=i}^{L} b(g)\right) \left(\frac{\beta}{\lambda + \beta}\right)^{i-\ell} \binom{i - \ell + r - 1}{r - 1}\right], \quad \ell = 1, \dots, L, \tag{2.2.5}$$

$$\theta(r+1) = \theta(r+2) = \dots = \theta(S) = G^{-1} \left(\frac{\lambda + \nu}{\lambda}\right)^{r}, \tag{2.2.6}$$

[7]published as [KD15a, Proposition 1 in Section 4]

where the normalization constant G is chosen such that

$$\sum_{k\in K}\theta(k)=1.$$

For $r=0$ we obtain $\theta=(\theta(k):k\in K)$ with normalization constant G as

$$\theta(0,\ell)=G^{-1}\frac{\lambda}{\beta}\left[\sum_{i=\ell}^{L}b(i)\right],\qquad \ell=1,\dots,L, \tag{2.2.7}$$

$$\theta(1)=\theta(2)=\dots=\theta(S)=G^{-1}. \tag{2.2.8}$$

Proof. The inventory management process under (r,S)-policy with distribution function B of the lead times fits into the definition of the environment process by setting

$$K=\{r+1,r+2,\dots,S\}\cup(\{0,1,\dots,r\}\times\{L,\dots,1\}),\quad K_B=\{0\}\times\{L,\dots,1\}.$$

The non-negative transition rates of (X,Y) are for $(n,k)\in E$

$$\begin{aligned}
q((n,k),(n+1,k))&=\lambda, && k\in K_W, n\geq 0,\\
q((n,k),(n-1,m))&=\mu(n)R(k,m), && k\in K_W, m\in K, n>0,\\
q((n,k),(n,m))&=v(k,m)\in\mathbb{R}_0^+, && k\neq m,\quad k,m\in K,\\
q((n,k),(i,m))&=0, && \text{otherwise for } (n,k)\neq(i,m)\in E;
\end{aligned}$$

where

$$\begin{aligned}
R(k,k-1)&=1 && \text{if } k\in\{r+2,...,S\},\\
R(r+1,(r,\ell))&=b(\ell) && \text{if } \ell\in\{L,\dots,1\},\\
R((j,\ell),(j-1,\ell))&=1 && \text{if } (j,\ell)\in\{1,\dots,r\}\times\{L,\dots,1\},\\
R(k,j)&=0 && \text{if } k,j\in K,\ \text{otherwise},
\end{aligned}$$

and

$$\begin{aligned}
v((j,\ell),(j,\ell-1))&=\beta && \text{if } j\in\{0,1,\dots,r\},\ell\in\{L,\dots,2\}\\
v((j,1),S)&=\beta && \text{if } j\in\{0,1,\dots,r\},\\
v(k,j)&=0 && \text{if } k,j\in K,\ \text{otherwise}.
\end{aligned}$$

Because $\lambda(n)=\lambda$ for all n and K is finite, Lemma 2.1.11 applies and we know that the steady state of the ergodic system is of product form

$$\pi(n,k)=C^{-1}\frac{\lambda^n}{\prod_{i=0}^{n-1}\mu(i+1)}\theta(k)\quad n.k)\in E. \tag{2.2.9}$$

We have to solve (2.1.18). By definition this is (with $R(k,k)=0$, $\forall k\in K\backslash\{0\}$, $R(0,0)=1$)

$$\begin{aligned}
&\theta(k)\left(1_{[k\in K_W]}\lambda+\sum_{m\in K\backslash\{k\}}v(k,m)\right)\\
&=\sum_{m\in K_W\backslash\{k\}}\theta(m)\left(\lambda(n)R(m,k)+v(m,k)\right)+\sum_{m\in K_B\backslash\{k\}}\theta(m)v(m,k).
\end{aligned} \tag{2.2.10}$$

(I) For $r > 0$, (2.2.10) translates into

$$\theta(S)\cdot\lambda = \sum_{j=0}^{r}\theta(j,1)\cdot\beta, \tag{2.2.11}$$

$$\theta(k)\cdot\lambda = \theta(k+1)\cdot\lambda, \qquad k = r+1,\dots,S-1, \tag{2.2.12}$$

$$\theta(r,\ell)\cdot(\lambda+\beta) = \theta(r+1)\cdot\lambda b(\ell) + \theta(r,\ell+1)\cdot\beta, \qquad 1\le\ell<L, \tag{2.2.13}$$

$$\theta(r,L)\cdot(\lambda+\beta) = \theta(r+1)\cdot\lambda b(L), \tag{2.2.14}$$

$$\theta(j,L)\cdot(\lambda+\beta) = \theta(j+1,L)\cdot\lambda, \qquad 1\le j<r, \tag{2.2.15}$$

$$\theta(j,\ell)\cdot(\lambda+\beta) = \theta(j+1,\ell)\cdot\lambda + \theta(j,\ell+1)\cdot\beta, \qquad 1\le j<r, 1\le\ell<L, \tag{2.2.16}$$

$$\theta(0,\ell)\cdot\beta = \theta(1,\ell)\cdot\lambda + \theta(0,\ell+1)\cdot\beta, \qquad 1\le\ell<L, \tag{2.2.17}$$

$$\theta(0,L)\cdot\beta = \theta(1,L)\cdot\lambda. \tag{2.2.18}$$

From (2.2.12) follows

$$\theta(S) = \theta(S-1) = \cdots = \theta(r+1), \tag{2.2.19}$$

and from (2.2.14) and (2.2.15) follows

$$\theta(j,L) = \theta(r+1)b(L)\left(\frac{\lambda}{\lambda+\beta}\right)^{r+1-j}. \tag{2.2.20}$$

From (2.2.20) (for $j = r$) and (2.2.13) follows directly

$$\theta(r,\ell) = \theta(r+1)\frac{\lambda}{\lambda+\beta}\sum_{i=\ell}^{L} b(i)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}, \quad 1\le\ell<L. \tag{2.2.21}$$

Up to now we obtained the expressions on border lines of the array $(\theta(j,\ell) : 1 \le j \le r, 1 \le \ell \le L)$ which can be filled step by step via (2.2.16). The proposed solution is

$$\theta(r-h,\ell) = \theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^{h+1}\sum_{i=\ell}^{L} b(i)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+h}{h}, \tag{2.2.22}$$

for $h = 0,1,\dots,r-1, \ell = 1,\dots,L$ fits with (2.2.21) ($h = 0$ with $\binom{i-l}{0} = 1$) and (2.2.20). Inserting (2.2.22) into (2.2.16) verifies (2.2.22) by a two-step induction with help by the elementary formula $\binom{a}{n} + \binom{a}{n-1} = \binom{a+1}{n}$.

For computing the residual boundary probabilities $\theta(0,\ell)$ we need some more effort. The proposed solution is for $\ell = 1,\dots,L$

$$\theta(0,\ell) = \theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^{r}\frac{\lambda}{\beta}\left[\sum_{i=\ell}^{L}\sum_{g=i}^{L} b(g)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-1}{r-1}\right]. \tag{2.2.23}$$

From (2.2.18) and (2.2.20) we obtain

$$\theta(0,L) = \theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^{r}\frac{\lambda}{\beta}b(L), \tag{2.2.24}$$

which fits into (2.2.23), and it remains to check the recursion (2.2.17). This amounts to compute

$$
\begin{aligned}
&\theta(1,\ell)\cdot\frac{\lambda}{\beta}+\theta(0,\ell+1)\\
&=\theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^r\sum_{i=\ell}^{L}b(i)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-1}{r-1}\cdot\frac{\lambda}{\beta}+\\
&\quad+\theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^r\frac{\lambda}{\beta}\left[\sum_{i=\ell+1}^{L}\sum_{g=i}^{L}b(g)\left(\frac{\beta}{\lambda+\beta}\right)^{i-(\ell+1)}\binom{i-(\ell+1)+r-1}{r-1}\right]\\
&=\theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^r\frac{\lambda}{\beta}\left[\sum_{i=\ell+1}^{L}\left\{\sum_{g=i}^{L}b(g)\left(\frac{\beta}{\lambda+\beta}\right)^{i-(\ell+1)}\binom{\overbrace{i-(\ell+1)}^{=(i-1)-\ell}+r-1}{r-1}\right.\right.\\
&\qquad\left.+b(i-1)\left(\frac{\beta}{\lambda+\beta}\right)^{(i-1)-\ell}\binom{(i-1)-\ell+r-1}{r-1}\right\}+\\
&\qquad\left.+b(L)\left(\frac{\beta}{\lambda+\beta}\right)^{L-\ell}\binom{L-\ell+r-1}{r-1}\right]\\
&=\theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^r\frac{\lambda}{\beta}\left[\sum_{i=\ell+1}^{L}\left\{\sum_{g=i-1}^{L}b(g)\left(\frac{\beta}{\lambda+\beta}\right)^{(i-1)-\ell}\binom{(i-1)-\ell+r-1}{r-1}\right\}\right.\\
&\qquad\left.+b(L)\left(\frac{\beta}{\lambda+\beta}\right)^{L-\ell}\binom{L-\ell+r-1}{r-1}\right]\\
&=\theta(r+1)\left(\frac{\lambda}{\lambda+\beta}\right)^r\frac{\lambda}{\beta}\left[\sum_{i=\ell}^{L-1}\left\{\sum_{g=i}^{L}b(g)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-1}{r-1}\right\}\right.\\
&\qquad\left.+b(L)\left(\frac{\beta}{\lambda+\beta}\right)^{L-\ell}\binom{L-\ell+r-1}{r-1}\right]=\theta(0,\ell).
\end{aligned}
$$

Setting

$$\theta(r+1)=G^{-1}\left(\frac{\lambda+\nu}{\lambda}\right)^r$$

completes the proof in case of $r=0$.

(II) For $r>0$, (2.2.10) translates into

$$
\begin{aligned}
\theta(S)\cdot\lambda &= \theta(0,1)\cdot\beta, && && (2.2.25)\\
\theta(k)\cdot\lambda &= \theta(k+1)\cdot\lambda, && k=1,\dots,S-1, && (2.2.26)\\
\theta(0,\ell)\cdot\beta &= \theta(1)\cdot\lambda\cdot b(L)+\theta(0,\ell+1)\cdot\beta, && 1\le\ell<L, && (2.2.27)\\
\theta(0,L)\cdot\beta &= \theta(1)\cdot\lambda\cdot b(L). && && (2.2.28)
\end{aligned}
$$

From (2.2.26) follows

$$\theta(S)=\theta(S-1)=\cdots=\theta(1), \tag{2.2.29}$$

and we will show that

$$\theta(0,\ell) = \theta(1) \cdot \left(\frac{\lambda}{\beta}\right) \left[\sum_{i=\ell}^{L} b(i)\right], \quad \ell = 1, \ldots, L, \tag{2.2.30}$$

holds. For $\ell = L$ this is immediate from (2.2.28), and for $\ell < L$ it follows by induction from (2.2.27). Setting $\theta(1) = G^{-1}$ completes the proof in case of $r = 0$. □

Remark 2.2.10. For $r > 0$ we can write (2.2.5) as

$$\theta(0,\ell) = G^{-1}\frac{\lambda}{\beta}\left[\sum_{i=\ell}^{L}\left(\sum_{g=i}^{L} b(g)\right)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-1}{i-\ell}\right],$$
$$\ell = 1, \ldots, L,$$

and can extend this formula to the case $r = 0$. This yields with $\binom{-1}{0} = 1$ explicitly

$$\theta(0,\ell) \quad = \quad G^{-1}\frac{\lambda}{\beta}\left[\sum_{i=\ell}^{L} b(i)\right], \quad \ell = 1, \ldots, L.$$

Corollary 2.2.11. [8]*In steady state the marginal probabilities for the inventory at hand have the following simple representation.*

Denote by ν^{-1} the expected lead time.

Let U denote a random variable distributed according to $b = (b(\ell) : 1 \le \ell \le L)$, and let U_e denote a random variable distributed according to the "equilibrium distribution" of U, resp. b, i,e.

$$P(U_e = i) = \frac{1}{E(U)}\sum_{g=i}^{L} b(g), \quad 1 \le i \le L.$$

Let $W(u,\alpha)$ denote a random variable distributed according to a negative binomial distribution $Nb^0(u,\alpha)$ with parameters $u \in \mathbb{N}$ and $\alpha \in (0,1)$, i.e.,

$$P(W(u,\alpha) = i) = \binom{i+u-1}{u-1}\alpha^u(1-\alpha)^i, \quad i \in \mathbb{N}.$$

Let I denote a random variable distributed according to the marginal steady-state probability for the inventory size. Then for $j = 1, \ldots, r$

$$P(I = j) = G^{-1}\left(\frac{\lambda+\beta}{\lambda}\right)^r \cdot P\left(W\left(r+1-j, \frac{\lambda}{\lambda+\beta}\right) < U\right), \tag{2.2.31}$$

and

$$P(I = 0) = G^{-1}\frac{\lambda}{\nu}\left(\frac{\lambda+\beta}{\lambda}\right)^r \cdot P\left(W\left(r, \frac{\lambda}{\lambda+\beta}\right) < U_e\right). \tag{2.2.32}$$

For $j = r+1, \ldots, S$ (2.2.6) applies directly:

$$P(I = r+1) = \cdots = P(I = S) = G^{-1}\left(\frac{\lambda+\nu}{\lambda}\right)^r.$$

[8]published as [KD15a, Corollary 2 in Section 4]

Proof. For $j = 1, \dots, r$ we have

$$P(I=j) = G^{-1}\left(\frac{\lambda+\beta}{\lambda}\right)^{j-1}\sum_{\ell=1}^{L}\sum_{i=\ell}^{L} b(i)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-j}{r-j}$$

$$= G^{-1}\left(\frac{\lambda+\beta}{\lambda}\right)^{j-1}\sum_{i=1}^{L} b(i)\sum_{\ell=1}^{i}\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-j}{r-j}$$

$$= G^{-1}\left(\frac{\lambda+\beta}{\lambda}\right)^{j-1}\sum_{i=1}^{L} b(i)\sum_{g=0}^{i-1}\left(\frac{\beta}{\lambda+\beta}\right)^{g}\binom{g+r-j}{r-j}$$

$$= G^{-1}\left(\frac{\lambda+\beta}{\lambda}\right)^{j-1}\left(\frac{\lambda+\beta}{\lambda}\right)^{r+1-j}$$

$$\sum_{i=1}^{L} b(i)\sum_{g=0}^{i-1}\binom{g+(r+1-j)-1}{(r+1-j)-1}\left(\frac{\lambda}{\lambda+\beta}\right)^{r+1-j}\left(\frac{\beta}{\lambda+\beta}\right)^{g}$$

$$= G^{-1}\left(\frac{\lambda+\beta}{\lambda}\right)^{r}\sum_{i=1}^{L} b(i)\cdot P(W(r+1-j,\frac{\lambda}{\lambda+\beta}) < i),$$

and for $j = 0$ we have

$$P(I=0) = G^{-1}\left(\frac{\lambda}{\beta}\right)\sum_{\ell=1}^{L}\sum_{i=\ell}^{L}\sum_{g=i}^{L} b(g)\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-1}{r-1}$$

$$= G^{-1}\left(\frac{\lambda}{\beta}\right)\sum_{i=1}^{L}\sum_{\ell=1}^{i}\left(\frac{\beta}{\lambda+\beta}\right)^{i-\ell}\binom{i-\ell+r-1}{r-1}\sum_{g=i}^{L} b(g)$$

$$= G^{-1}\underbrace{\left(\frac{\lambda}{\beta}\cdot E(V)\right)}_{=\lambda/\nu}\left(\frac{\lambda+\beta}{\lambda}\right)^{r}$$

$$\cdot\sum_{i=1}^{L}\underbrace{\left(\frac{1}{E(U)}\sum_{g=i}^{L} b(g)\right)}_{=:P(V_e=i)}\sum_{f=0}^{i-1}\binom{f+r-1}{r-1}\left(\frac{\lambda}{\lambda+\beta}\right)^{r}\left(\frac{\beta}{\lambda+\beta}\right)^{f}$$

$$= G^{-1}\left(\frac{\lambda}{\nu}\right)\left(\frac{\lambda+\beta}{\lambda}\right)^{r} P\left(W\left(r+1-1,\frac{\lambda}{\lambda+\beta}\right) < U_e\right).$$

□

We now revisit the results from [SHH11] and [SAH13] for queueing-inventory systems under (r,Q) policy. We allow additionally the service rate of the server to depend on the queue length of the system. We assume that the lead time distribution is of phase type.

We enlarge the phase space of the system, i.e. the state space of the inventory process Y. Whenever there is an ongoing lead time, i.e., when inventory at hand is less than $r+1$, we count the number of residual successive i.i.d. $\exp(\beta)$-distributed lead time phases which must expire until the replenishment arrives at the inventory.

The state space of (X, Y) then is $E = \mathbb{N}_0 \times K$ with

$$K = \{r+1, r+2, \ldots, r+Q\} \cup (\{0, 1, \ldots, r\} \times \{L, \ldots, 1\}),$$

and (X, Y) is irreducible on E.

Proposition 2.2.12. [9]*$M/M/1/\infty$ system with (r, Q)-policy, phase-type replenishment lead time, state dependent service rates $\mu(n)$, and lost sales.*

The lead time distribution has a distribution function B from (2.2.2). *We assume that (X, Y) is positive recurrent and denote its steady-state distribution by*

$$\pi = (\pi(n, k) : n \in \mathbb{N}_0 \times K).$$

The steady-state distribution π of (X, Y) is of product form. With normalization constant C

$$\pi(n, k) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda}{\mu(i+1)} \cdot \theta(k), \tag{2.2.33}$$

where $\theta = (\theta(k) : k \in K)$ can be obtained from formula (3) in [SAH13], and the subsequent formulas (4) - (10) there.

Proof. The proof is in its first part similar to that of Proposition 2.2.9 because the inventory management process under (r, Q)-policy with distribution function B of the lead times fits into the definition of the environment process by setting

$$K = \{r+1, r+2, \ldots, r+Q\} \cup (\{0, 1, \ldots, r\} \times \{L, \ldots, 1\}), \quad K_B = \{0\} \times \{L, \ldots, 1\}.$$

Because $\lambda(n) = \lambda$ for all n and K is finite, Lemma 2.1.11 applies and we know that the steady state of the ergodic system is of product form

$$\pi(n, k) = C^{-1} \frac{\lambda^n}{\prod_{i=0}^{n-1} \mu(i+1)} \cdot \theta(k), \quad (n, k) \in E. \tag{2.2.34}$$

Thus the product form statement is proven with the required marginal queue length distribution.

In a second part we have to compute the $\theta(k)$ which is to solve (2.1.18). This equation is independent of n, especially independent of the $\mu(n)$.

Therefore the solution in the case of state independent service rates ($\mu(n) \to \mu$) from [SAH13] must be the solution in the present slightly more general setting as well. □

2.2.2. Unreliable servers

In [SD03a] networks of queues with unreliable servers were investigated which admit product form steady states in twofold way: The joint queue length vector of the system (which in general is not a Markov process) is of classical product form as in Jackson's Theorem and the availability status of the nodes as a set valued supplementary variable process constitutes an additional product factor attached to the joint queue length vector.

We show for the case of a single server which is unreliable and breaks down due to influences from an environment that a similar product form result follows from our Theorem 2.1.5. We allow for a much more complicated breakdown and repair process as that investigated in [SD03a].

[9]published as [KD15a, Proposition 2 in Section 4][KD15a, Proposition 2 in Section 4]

Example 2.2.13. [10]There is a single exponential server with with Poisson-λ arrival stream and state dependent service rates $\mu(n)$. The server acts in a random environment which changes over time. The server breaks down with rates depending on the state of the environment and is repaired after a breakdown with repair intensity depending on the state of the environment as well. Whenever the server is broken down, new arrivals are not admitted and are lost to the system forever.

The system is described by a two-dimensional Markov process $(X, Y) = ((X(t), Y(t)) : t \in [0, \infty))$ with state space $E = \mathbb{N}_0 \times K$. K is the (countable) environment space of the process, whereas $\mathbb{N}_0$ denotes the queue length. (X, Y) is assumed to be irreducible.

The environment space of the process is partitioned into disjoint nonempty components $K := K_W \uplus K_B$, and whenever Y enters K_B the server breaks down immediately, and will be repaired when Y enters K_W again.

The non-negative transition rates of (X, Y) are for $(n, k) \in E$

$$\begin{aligned} q((n,k),(n+1,k)) &= \lambda, && k \in K_W, \\ q((n,k),(n-1,m)) &= \mu(n)R(k,m), && k \in K_W, n > 0, \\ q((n,k),(n,m)) &= v(k,m) \in \mathbb{R}_0^+, && k \neq m, \\ q((n,k),(l,m)) &= 0, && \text{otherwise for } (n,k) \neq (l,m), \end{aligned} \tag{2.2.35}$$

and from Corollary 2.1.14 we directly obtain in case of ergodicity the product form steady-state distribution if the set K is finite. For an infinite set K, we additionally require summability condition (2.1.17) and that the equation (2.1.18) has a unique positive stochastic solution. Then we get the unique product form distribution according to Corollary 2.1.10.

An interesting observation is, that we can model general distributions for the successive times the system is functioning and similarly for the repair times.

By suitably selected structures for the $v(\cdot, \cdot)$ we can incorporate dependent up and down times.

The distinctive feature which sets the difference to the breakdown mechanism in [SD03a] is that breakdowns can be directly connected with expiring service times via the stochastic matrix R, which is visible from (2.2.35). This widens applicability of the mechanism considerably.

2.2.3. Tandem system with finite intermediate buffer

Modeling[11] multi-stage production lines by serial tandem queues is standard technique. In the simplest case with Poisson arrivals and with exponential production times for one unit in each stage the model fits into the realm of Jackson network models as long as the buffers between the stages have infinite capacity. Consequently, ergodic systems under this modeling approach have a product form steady-state distribution.

With respect to steady-state analysis the picture changes completely if the buffers between the stages have only finite capacity, no simple solutions are available. Direct

[10]This is an improved version of [KD15a, Example 9]. We explicitly point out here the summability condition (2.1.17) and existence of a unique positive stochastic solution of equation (2.1.18) if the set K is infinite.

[11]published as [KD15a, Subsection 4.2.2]

numerical analysis or simulations are needed, or we have to resort to approximations. A common procedure is to use product form approximations which are developed by decomposition methods. A survey on general networks with blocking is [BDO01], special emphasis to modeling manufacturing flow lines is given in the survey [DG92].

The same class of problems and solutions are well known in teletraffic networks where finite buffers are encountered, for surveys see [Onv90] and [Per90].

A systematic study of how to use product form networks as upper or lower bounds (in a specified performance metric) is given in [Dij93]. A closed 3-station model which is related to the one given below is discussed in [Dij93, Section 4.5.1], where product form lower and upper bounds are proposed.

Van Dijk [Dij11, p.44] describes a tandem system with $\mu(n) = \mu$, $\nu(k) = \nu$ which leads to a product form. We extend this model by allowing more general service rates $\mu(n)$ and ν_k.

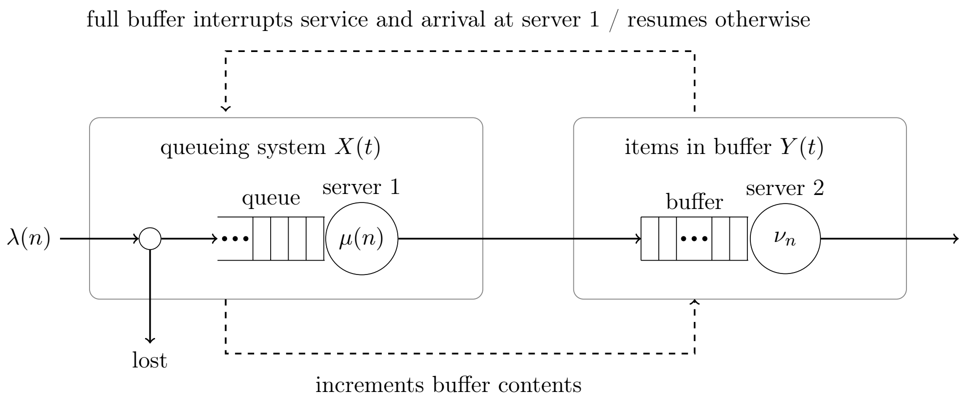


Figure 2.2.3.: Tandem system with finite intermediate buffer of size N.

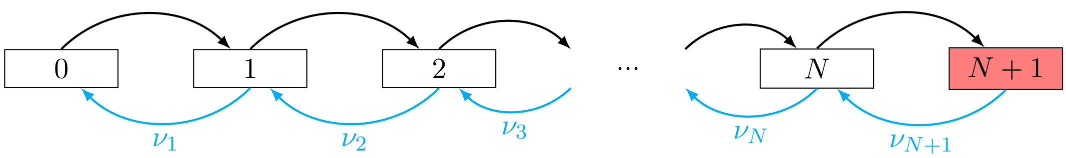


Figure 2.2.4.: Environment transition and interaction diagram of $M/M/1/\infty$ tandem system with finite intermediate buffer of size N.

We consider a two-stage single server tandem queueing system where the first station has ample waiting space while the buffer can contain maximal N waiting units, $N < \infty$, i.e. there can at most $N + 1$ units be stored in the system which have been processed at the first stage. It follows that for the system must be determined a blocking regime, which enforces the first station to stop production when the intermediate buffer reaches its capacity of N waiting units. We apply the **blocking-before-service** regime [Per90, p. 455]: Whenever the second station is full, the server at the first station does not start serving the next customer. When a departure occurs from the second station, the first station is unblocked immediately and resumes its service. Additionally, we require that the first station, when blocked, does not accept new customers, i.e., it is completely blocked.

The arrival stream is Poisson-λ, service rates are state dependent with $\mu(n)$ at the first station and ν_k at the second. To emphasize the modeling of the second server as an environment we use the notation ν_k instead of well known from the literature notation $\nu(k)$ for the service rate of the second server with k present customers. The standard independence assumption are assumed to hold, service at both stations is on FCFS basis.

This makes the joint queue length process (X, Y) Markov with state space $E := \mathbb{N}_0 \times \{0, 1, \dots, N, N+1\}$.

The non-negative transition rates are

$$
\begin{aligned}
q((n,k),(n+1,k)) &= \lambda, && k \le N,\\
q((n,k),(n-1,k+1)) &= \mu(n), && n > 0, k \le N,\\
q((n,k),(n,k-1)) &= \nu_k, && 1 \le k \le N+1,\\
q((n,k),(j,m)) &= 0, && \text{otherwise for } (n,k) \ne (j,m).
\end{aligned}
$$

We fit this model into the formalism of Section 2.1.1 by setting

$$
\begin{aligned}
K &= \{0,1,...,N+1\}, & K_B &= \{N+1\},\\
R(k,k+1) &= 1, 0 \le k \le N, & R(N+1,N+1) &= 1,
\end{aligned}
$$

$$
v(k,m) = \begin{cases} \nu_k, & \text{if } 1 \le k \le N+1, \text{ and } m = k-1,\\ & 0, \text{otherwise for } k \ne m. \end{cases}
$$

From Corollary 2.1.13 we conclude that for the ergodic process (X, Y) the steady-state distribution has product form

$$
\pi(n,k) = C^{-1} \frac{\lambda^n}{\prod_{i=0}^{n-1} \mu(i+1)} \theta(k), \qquad (n,k) \in E, \tag{2.2.36}
$$

with probability distribution θ on K and normalization constant

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

It remains to determine θ from Corollary 2.1.10, which is (2.1.18).

So, the Q_{red} matrix is explicitly

$$
Q_{\text{red}} = \left(\begin{array}{c|cccccc}
 & 0 & 1 & 2 & & N & N+1 \\
\hline
0 & -\lambda & \lambda & & & & \\
1 & \nu_1 & -(\nu_1+\lambda) & \lambda & & & \\
2 & & \nu_2 & \ddots & \ddots & & \\
 & & & \ddots & \ddots & \ddots & \\
N & & & & \ddots & -(\nu_N+\lambda) & \lambda \\
N+1 & & & & & \nu_{N+1} & -\nu_{N+1}
\end{array} \right).
$$

This is exactly the transition rate matrix of an $M/M/1/N$ queue with Poisson-λ arrivals and service rates ν_k and we have immediately

$$\theta(k) = G^{-1} \prod_{h=1}^{k} \frac{\lambda}{\nu_h}, \qquad 0 \leq k \leq N+1, \tag{2.2.37}$$

with normalization constant

$$G = \sum_{h=0}^{N+1} \left(\prod_{h=1}^{k} \frac{\lambda}{\nu_h} \right).$$

Remark 2.2.14. The result

$$\pi(n,k) = C^{-1} \prod_{i=0}^{n-1} \frac{\lambda^n}{\mu(i+1)} \cdot G^{-1} \prod_{h=1}^{k} \frac{\lambda}{\nu_h}, \quad (n,k) \in E,$$

is surprising, because it looks like an independence result with marginal distributions of two queues fed by Poisson-λ streams. Due to the interruptions, neither the arrival process at the first station nor the departure stream from the first node, which is the arrival stream to the second, is Poisson-λ. There seems to be no intuitive explanation of the results.

2.2.4. Unreliable $M/M/1/\infty$ queueing system with control of repair and maintenance[12]

In her PhD-thesis [Sau06, Section 3.2] Cornellia Sauer introduced degradable networks where failure behaviour was coupled with a service "counter". The counter is a special environment variable which is decreased right after a service and can be reseted by a repair or a preventive maintenance depending on its current value. Sauer discussed in Remark 3.2.8 there similarities between degradable network models with service counter and networks with inventories.

In this section we will analyze a queueing system, which utilizes a counter to control repair and preventive maintenance and for modeling of failure behaviour.

We consider a queueing system, where the server wears down during service. As a consequence the failure probability can change. We do not require the failure probability to increase. In some systems it is observed that whenever the server survives an initial period its reliability stays constant or even increases.

When the system breaks down it is repaired and thereafter resumes work as good as new. Furthermore, to prevent break downs, the system will be maintained after a prescribed (maximal) number N of services since the most recent repair or maintenance. During repair or maintenance the system is blocked, i.e., no service is provided and no new job may join the system. These rejected jobs are *lost* to the system.

Subject to optimization is N – the maximal number of services, after which the system needs to be maintained.

[12] The content of this section is published in [KD15b].

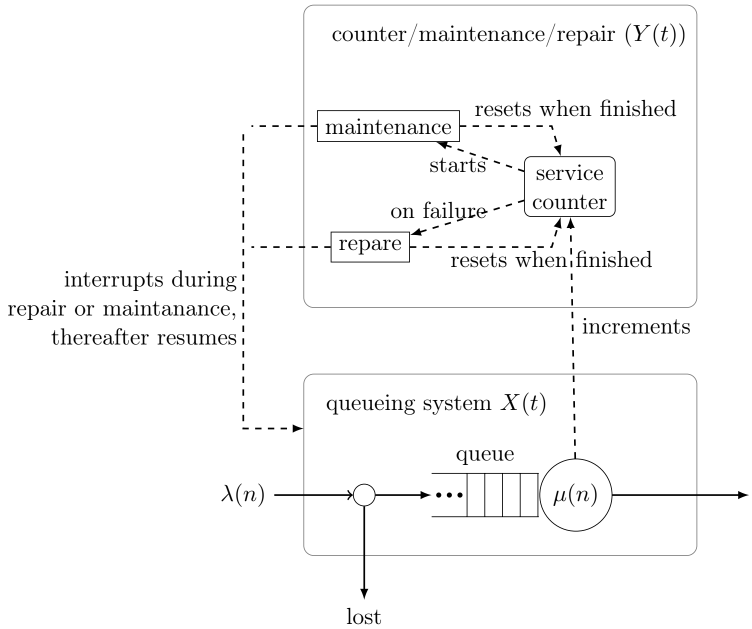


Figure 2.2.5.: $M/M/1/\infty$ unreliable loss system.

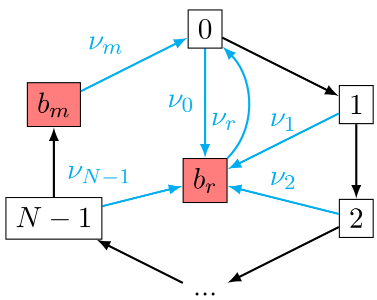


Figure 2.2.6.: Environment transition and interaction diagram for the $M/M/1/\infty$ unreliable system. The environment describes the service counter, repair state and maintenance state.

2.2.4.1. Model

We consider a production system which is modeled as an $M/M/1/\infty$ loss system. That is with λ Poisson input rate, service rates $\vec{\mu} := (\mu(n) : n \in \mathbb{N})$ depending on the number of customers in the system, $FCFS$ service regime, and environment states $K = K_W + K_B$.

The state space of the system is $E = \mathbb{N}_0 \times (\{0, 1, \ldots, N-1\} \cup \{b_m, \ldots, b_r\})$. The environment states $K_W = \{0, 1, \ldots, N-1\}$ indicate the number of services completed since the last repair or maintenance (service counter). N is the maximal number of services before maintenance is required. The additional environment states $K_B = \{b_m, b_r\}$ indicate when there is an ongoing maintenance (b_m) or repair (b_r).

We define a stochastic matrix $R \in [0,1]^{K \times K}$ which determines the behaviour of the "service counter". Transition rates $R(k, k+1) = 1$ for $0 \leq k \leq N-2$ govern counter increment and transition rate $R(N-1, b_m) = 1$ enforces mandatory maintenance after N services.

We use infinitesimal generator $V \in \mathbb{R}^{K \times K}$ to control failure, maintenance, and repair rates: $v(k, b_r) = \nu_k$ are failure rates after k complete services, $v(b_m, 0) = \nu_m$ is the maintenance rate and $v(b_r, 0) = \nu_r$ is the repair rate.

We define by $Z = (X, Y)$ the joint queue length and environment process of this system and make the usual independence assumptions for the queue and the environment. The Z is Markov process, which we assume to be ergodic.

The total costs of the system is determined by specific cost constants per unit of time: maintenance costs c_m, repair costs c_r, costs of non-availability c_b, and waiting costs in queue and in service per customer c_w. Therefore the cost function per unit of time in the respective states is

$$
\begin{aligned}
f : \mathbb{N}_0 \times K &\longrightarrow \mathbb{R} \\
f(n,k) &= \begin{cases} c_w \cdot n + c_b + c_m, & k = b_m, \\ c_w \cdot n + c_b + c_r, & k = b_r, \\ c_w \cdot n, & k \in K \backslash \{b_m, b_r\}. \end{cases}
\end{aligned}
$$

Our aim is to analyze the long-run system behaviour and to minimize the long-run average costs.

Proposition 2.2.15. [13] *The steady-state distribution of the system described above is*

$$\lim_{t \to \infty} P(X(t) = n, Y(t) = k) =: \pi(n,k) = \xi(n)\theta(k), \text{ with}$$

$$\xi(n) = \prod_{i=1}^{n} \frac{\lambda}{\mu(i)} \xi(0), \text{ and}$$

$$
\begin{aligned}
\theta(k) &= \prod_{i=1}^{k} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \theta(0), \qquad 0 \leq k \leq N-1, \\
\theta(b_m) &= \frac{\lambda}{\nu_m} \theta(N-1) = \frac{\lambda}{\nu_m} \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \theta(0),
\end{aligned}
$$

[13] published as [KD15b, Theorem 1]

$$\theta(b_r) = \left(\frac{(\nu_0 + \lambda)}{\nu_r} - \frac{\lambda}{\nu_r} \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \right) \theta(0),$$

$$\theta(0) = \frac{1}{\left(\frac{(\nu_0+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r} \right) \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i+\lambda} \right)^i \right) + \sum_{k=0}^{N-1} \prod_{i=1}^{k} \left(\frac{\lambda}{\nu_i+\lambda} \right)^i}.$$

Proof. The $M/M/1/\infty$ system with unreliable server, maintenance and repair is a loss system in a random environment with

$$V =$$

$$\left(\begin{array}{c|ccccccccc}
 & 0 & 1 & 2 & \dots & N-3 & N-2 & N-1 & b_m & b_r \\
\hline
0 & -\nu_0 & 0 & 0 & & 0 & 0 & 0 & 0 & \nu_0 \\
1 & 0 & -\nu_1 & 0 & & 0 & 0 & 0 & 0 & \nu_1 \\
2 & 0 & 0 & -\nu_2 & & 0 & 0 & 0 & 0 & \nu_2 \\
\vdots & & & & \ddots & & & & & \vdots \\
N-3 & 0 & 0 & 0 & \dots & -\nu_{N-3} & 0 & 0 & 0 & \nu_{N-3} \\
N-2 & 0 & 0 & 0 & \dots & 0 & -\nu_{N-2} & 0 & 0 & \nu_{N-2} \\
N-1 & 0 & 0 & 0 & \dots & 0 & 0 & -\nu_{N-1} & & \nu_{N-1} \\
b_m & \nu_m & 0 & 0 & \dots & 0 & 0 & & -\nu_m & 0 \\
b_r & \nu_r & 0 & 0 & \dots & 0 & 0 & & 0 & -\nu_r
\end{array} \right)$$

and

$$R =$$

$$\left(\begin{array}{c|ccccccccc}
 & 0 & 1 & 2 & \dots & N-3 & N-2 & N-1 & b_m & b_b \\
\hline
0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\
\vdots & & & & & \ddots & & & & \vdots \\
N-3 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 \\
N-2 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \\
N-1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 \\
b_m & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 \\
b_r & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 1
\end{array} \right).$$

The matrix Q_{red} from Corollary 2.1.10 is

$$Q_{\text{red}} = \lambda I_W (R - I) + V =$$

$$\left(\begin{array}{c|ccccccccc}
 & 0 & 1 & 2 & \dots & N-3 & N-2 & N-1 & b_m & b_r \\
\hline
0 & -\nu_0 - \lambda & \lambda & 0 & & 0 & 0 & 0 & 0 & \nu_0 \\
1 & 0 & -\nu_1 - \lambda & \lambda & & 0 & 0 & 0 & 0 & \nu_1 \\
2 & 0 & 0 & -\nu_2 & \ddots & 0 & 0 & 0 & 0 & \nu_2 \\
\vdots & & & & \ddots & \ddots & & & & \vdots \\
N-3 & 0 & 0 & 0 & \dots & -\nu_{N-3} - \lambda & \lambda & 0 & 0 & \nu_{N-3} \\
N-2 & 0 & 0 & 0 & \dots & 0 & -\nu_{N-2} - \lambda & \lambda & 0 & \nu_{N-2} \\
N-1 & 0 & 0 & 0 & \dots & 0 & 0 & -\nu_{N-1} - \lambda & \lambda & \nu_{N-1} \\
b_m & \nu_m & 0 & 0 & \dots & 0 & 0 & 0 & -\nu_m & 0 \\
b_r & \nu_r & 0 & 0 & \dots & 0 & 0 & 0 & 0 & -\nu_r
\end{array} \right)$$

and the steady-state solution of the system has a product form with marginal distribution θ solving $\theta Q_{\text{red}} = 0$.

We now solve the equation $\theta Q_{\text{red}} = 0$.

For $k \in \{1, 2, \dots, N-1\}$ it follows directly

$$\begin{aligned}
\lambda\theta(k-1) - (\nu_k + \lambda)\theta(k) = 0 \\
\implies \theta(k) &= \frac{\lambda}{\nu_k + \lambda} \theta(k-1) \\
\implies \theta(k) &= \prod_{i=1}^{k} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \theta(0).
\end{aligned}$$

For $k = b_m$ we have

$$\lambda\theta(N-1) - \nu_m\theta(b_m) = 0$$
$$\theta(b_m) = \frac{\lambda}{\nu_m}\theta(N-1) = \frac{\lambda}{\nu_m}\prod_{i=1}^{N-1}\left(\frac{\lambda}{\nu_i+\lambda}\right)^i\theta(0).$$

Finally, for $k = 0$ we obtain

$$-(\nu_0+\lambda)\theta(0) + \nu_m\theta(b_m) + \nu_r\theta(b_r) = 0$$
$$\implies \theta(b_r) = \frac{(\nu_0+\lambda)}{\nu_r}\theta(0) - \frac{\nu_m}{\nu_r}\theta(b_m)$$
$$= \left((\nu_0+\lambda) - \lambda\prod_{i=1}^{N-1}\left(\frac{\lambda}{\nu_i+\lambda}\right)^i\right)\frac{1}{\nu_r}\theta(0).$$

This leads to

$$\theta(0) = \frac{1}{\left(\frac{(\nu_0+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right)\prod_{i=1}^{N-1}\left(\frac{\lambda}{\nu_i+\lambda}\right)^i\right) + \sum_{k=0}^{N-1}\prod_{i=1}^{k}\left(\frac{\lambda}{\nu_i+\lambda}\right)^i}.$$

□

2.2.4.2. Average costs

We will analyze average long term costs of the system with different N - the maximal number of services, after which system needs to be maintained. In order to distinguish the steady-state distribution of the systems with different parameters N we will denote them π_N and θ_N.

Lemma 2.2.16. *The optimal solution for the problem described in Section 2.2.4.1 is*

$$\arg\min(g(N))$$

with

$$g(N) := (c_b + c_m)\,\theta_N(b_m) + (c_b + c_r)\,\theta_N(b_r).$$

Proof. Due to ergodicity it holds

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T f\left(X(\omega)_t, Y_t(\omega)\right)dt = \sum_{(n,k)} f(n,k)\pi_N(n,k) =: \bar{f}(N), \qquad P\text{-a.s.}$$

Using product form properties of the system we get

$$\bar{f}(N) = (c_b + c_m)\,\theta_N(b_m) + (c_b + c_r)\,\theta_N(b_r) + \underbrace{c_w\sum_{n=1}^{\infty} n\xi(n)}_{\text{independent of } N}$$

$$\implies \arg\min(\bar{f}(N)) = \arg\min(g(N)).$$

□

Example 2.2.17. We consider two unreliable systems with parameters $\lambda = 1$, $\mu = 1.5$, $\nu_m = 0.3$, $\nu_r = 0.1$, costs $c_m = 1$, $c_r = 2$, $c_b = 1$ and different linear increasing functions ν_k determining wearout.

In case $\nu_k = 0.01k$, the optimal number of services after which maintenance should be performed $N = \arg\min(g(N)) = 6$. See Figure 2.2.7a on page 65.

In the case $\nu_k = 0.001k$ the optimal value is $N = 23$. See Figure 2.2.7b on page 65.

The case of constant failure rate ν (independent of k, i. e., of no aging) is often of particular interest. For example it is common practice to substitute complex varying system parameters by average values. In the following corollary and the remark we will investigate the properties of such a system in detail.

Corollary 2.2.18. *If ν is constant the function $g(N)$ is either strictly monotone increasing, or strictly monotone decreasing, or constant. For each type of these monotonicities there exists parameters which generate this type.*

Proof. We analyze the difference

$$g(N+1) - g(N) = (c_b + c_m)\left(\theta_{N+1}(b_m) - \theta_N(b_m)\right) + (c_b + c_r)\left(\theta_{N+1}(b_r) - \theta_N(b_r)\right).$$

We calculate $\theta_N(0)$ for constant $\nu_i \equiv \nu$, we will use $a := \left(\frac{\lambda}{\nu+\lambda}\right)$ to simplify the notation

$$\begin{aligned}
\theta_N(0) &= \frac{1}{\left(\frac{(\nu+\lambda)}{\nu_r} - \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right) a^{N-1}\right) + \sum_{k=0}^{N-1} a^k} \\
&= \frac{1}{\left(\frac{(\nu+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right) a^{N-1}\right) + \frac{(1-a^N)}{(1-a)}} \\
&= \frac{(1-a)}{\left(\frac{(v+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right) a^{N-1}\right)(1-a) + (1-a^N)}.
\end{aligned}$$

Using the result above we calculate $\theta_N(b_m)$

$$\begin{aligned}
\theta_N(b_m) &= \frac{\frac{\lambda}{\nu_m}\left(\frac{\lambda}{\nu+\lambda}\right)^{N-1}(1-a)}{\left(\frac{(v+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right) a^{N-1}\right)(1-a) + (1-a^N)} \\
&= \frac{\overbrace{\left(\frac{\lambda}{\nu_m} a^{N-1}\right)(1-a)}^{=:Nom(N,b_m)}}{\underbrace{\left(\frac{(v_0+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right) a^{N-1}\right)(1-a) + \left(1-a^N\right)}_{=:Den(N)}} \\
&= \frac{Nom(N, b_m)}{Den(N)},
\end{aligned}$$

$$\theta_{N+1}(b_m) - \theta_N(b_m) = \frac{Nom(N+1, b_m)Den(N) - Nom(N, b_m)Den(N+1)}{Den(N+1)Den(N)}.$$

The common denominator of the difference $Den(N+1)Den(N)$ is positive, therefore we focus on nominator $Nom(N+1, b_m)Den(N) - Nom(N, b_m)Den(N+1) =$

$$= \frac{\lambda(1-a)a^{N-1}}{\nu_m \nu_r} \left(-a^2\nu + 2a\nu - \nu + a\nu_r - v_r - \lambda a^2 + 2\lambda a - \lambda\right)$$

Similarly we analyze the sign of the difference $\theta_{N+1}(b_r) - \theta_N(b_r)$:

$$\begin{aligned}
\theta_N(b_r) &= \left(\frac{(\nu+\lambda)}{\nu_r} - \frac{\lambda}{\nu_r}\left(\frac{\lambda}{\nu+\lambda}\right)^{N-1}\right)\theta_N(0) \\
&= \frac{\left(\frac{(v+\lambda)}{\nu_r} - \frac{\lambda}{\nu_r}a^{N-1}\right)}{\left(\frac{(\nu+\lambda)}{\nu_r} - \frac{\lambda}{\nu_r}a^{N-1}\right) + \frac{\lambda}{\nu_m}a^{N-1} + \sum_{k=0}^{N-1} a^k} \\
&= \frac{\overbrace{\left(\frac{(v+\lambda)}{\nu_r} - \frac{\lambda}{\nu_r}a^{N-1}\right)(1-a)}^{=:Nom(N,b_r)}}{\underbrace{\left(\frac{(v+\lambda)}{\nu_r} + \left(\frac{\lambda}{\nu_m} - \frac{\lambda}{\nu_r}\right)a^{N-1}\right)(1-a) + \left(1-a^N\right)}_{=:Den(N)}}.
\end{aligned}$$

$$\theta_{N+1}(b_r) - \theta_N(b_r) = \frac{Nom(N+1, b_r)Den(N) - Nom(N, b_r)Den(N+1)}{Den(N+1)Den(N)}$$

with $Nom(N+1, b_r)Den(N) - Nom(N, b_r)Den(N+1) =$

$$\frac{(1-a)a^{N-1}}{\nu_m\nu_r}\left(a^2\nu_m\nu - a\nu_m\nu + \lambda a^2\nu - 2\lambda a\nu + \lambda\nu + \lambda a^2\nu_m - 2\lambda a\nu_m + \lambda\nu_m + \lambda^2 a^2 - 2\lambda^2 a + \lambda^2\right)$$

and therefore

$$\begin{aligned}
&g(N+1) - g(N) \\
&= \frac{(1-a)a^{N-1}}{\nu_m\nu_r Den(N+1)Den(N)} \cdot \\
&\Big[(c_b + c_m)\lambda(-a^2\nu + 2a\nu - \nu + a\nu_r - \nu_r - \lambda a^2 + 2\lambda a - \lambda) \\
&\quad + (c_b + c_r) \cdot (a^2\nu_m\nu - a\nu_m\nu + \lambda a^2\nu - 2\lambda a\nu + \lambda\nu \\
&\qquad\qquad + \lambda a^2\nu_m - 2\lambda a\nu_m + \lambda\nu_m + \lambda^2 a^2 - 2\lambda^2 a + \lambda^2)\Big].
\end{aligned} \tag{2.2.38}$$

The sign of the difference depends only on the expression in the square brackets which is independent of N.

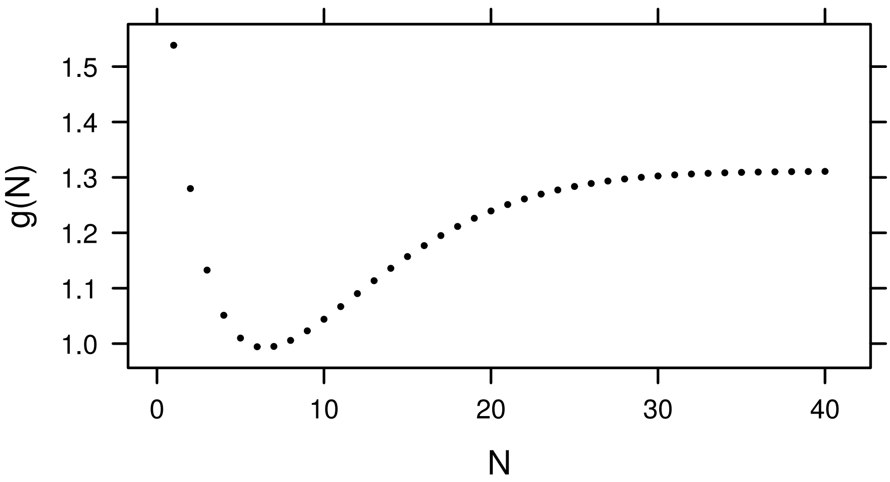
The results can be explained by the memoryless property of the exponential failure time distribution: If the failure rate is constant, the system behaviour stays the same, no matter how much time elapsed (or how many services are completed) since the last maintenance.

There exist parameters of the system such that the expression in squared brackets in (2.2.38) is negative, positive or zero. Consider for example $\lambda = 1$, $\nu = 0.04$, $\nu_m = 0.3$, $\nu_r = 0.1$, $c_m = 1$, and $c_b = 1$. For $c_r = 2$ this expression is negative, for $c_r = 10$ this expression is positive. According to Bolzano's theorem there exist $2 < c_r < 10$ such that the value of the polynomial function in square brackets is zero. □

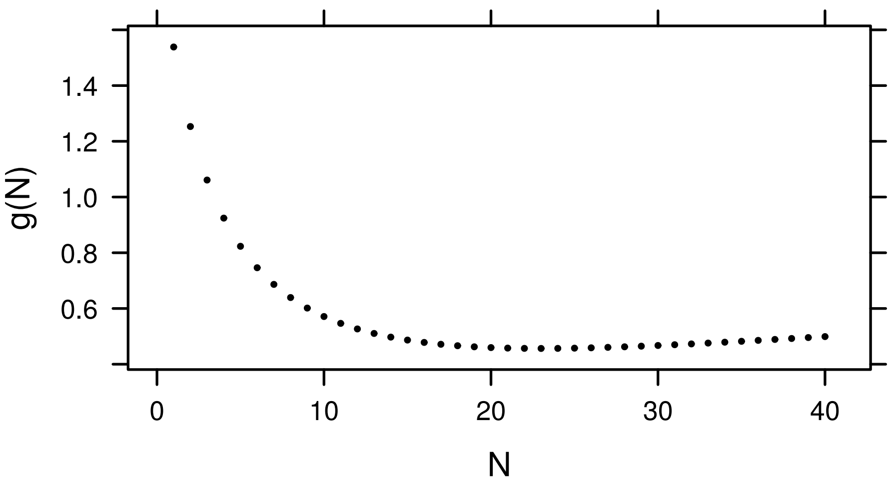
Remark 2.2.19. The most important consequence of Corollary 2.2.18 is that the costs function determines only three type of solutions:

- Maintain after the first service (cost function is strictly monotone increasing).
- Never maintain (cost function is strictly monotone decreasing).
- Maintain at any time (cost function stays constant).

As a consequence we see that models with constant failure rate are not well suited for drawing conclusions about, e.g., models with linear failure rates, see Example 2.2.17.



(a) $\nu_k = 0.01k$



(b) $\nu_k = 0.001k$

Figure 2.2.7.: Function $g(N)$ for linear break down rates ν_k in Example 2.2.17.

2.2.5. Modeling and performance analysis of a node in fault tolerant wireless sensor networks

The content of this section is published in [KD14]. The final publication is available at link.springer.com.

2.2.5.1. Introduction

Modeling fault tolerant (disruption tolerant) wireless sensor networks (DSN) is a challenging task due to the specific constraints imposed on the network structure and the principles the nodes have to follow to maintain connectedness of the network. The complexity of the models increases furthermore if the nodes are mobile and if energy efficiency is required. A typical way to resolve the latter problem is to reduce energy consumption by laying a node in sleep status whenever this is possible. In sleep status all activities of the node are either completely or almost completely interrupted. Clearly, this will have implications for availability of connections in the network.

In active mode the node undertakes several activities: Gathering data and putting the resulting data packets into its queue, receiving packets from other nodes which are placed in its queue (and relaying these packets when they arrive at the head of the node's queue), and processing the packets in the queue. Mobility requires routing decisions and routing evaluation which is connected with localization procedures.

Analytical performance analysis of DSN found in the literature usually follows a two-step procedure. (1) Investigate a single ("referenced") node, and (2) combine by some approximation procedure the interacting nodes to a separable network, for a review see [WDW07]. More detailed study of a specific node model is [Li11], other typical examples for the two-step procedure are [LTL05], [ZL11].

Our study is in part motivated by research in [WWDL07]. The protocol for the pervasive information gathering and processing described there (for more details on this protocol see [WWDL07]) consists basically on two "key components", (i) for data transmission, governed by "nodal delivery probabilities", and (ii) for queue management, governed by "message fault tolerances". Both of these components are complex and their interaction is a challenging problem for any modeling procedure.

[WWDL07] proceed as follows: In a first part "an overview of the dynamic delay/fault-tolerant mobile sensor network data delivery scheme" is provided with detailed recipes for the update procedure of the "nodal delivery probability" and the "message fault tolerance". In a second part complexity even of the single node's model is drastically reduced by not including the dynamics of these characteristics into the detailed analytical model. The authors argue: "While it is desirable to accurately analyze the data delivery scheme…, this is not practical given its complexity in data transmission and queue management" [WWDL07, p. 3290]. Consequently, for characterizing the behavior of a sensor node, (i) they fix for the node the nodal delivery probability as a constant depending mainly on the number of other nodes in its one-hop neighborhood, and (ii) the message fault tolerances are set to constant = 1, which means that no copy of a sent out message is kept by the sending node.

The aim of our presentation is to show that at least modeling of a single node and its **bidirectional interaction with the network**, can be done much more detailed than described in the cited literature, while still upcoming with closed formulas for the steady-state behavior and important structural theorems for the interaction between node and the network in equilibrium. An main observation will be that it is possible to reduce the complexity of the interacting processes in a similar way by **separability of the steady state** as complexity is reduced in the celebrated Jackson and Gordon-Newell networks.

Our result will not rely on a specific version of a DSN setting, and we will explain how to adapt our procedure with the selected model from [WWDL07] to other settings. We believe that the principles of our modeling procedure are rather generally to apply. Our procedure is:
Starting from the detailed description of a dynamic disruption/fault-tolerant mobile sensor node's data delivery scheme in [WWDL07], we develop an analytical model of a single node, the "referenced node" (RN) in the spirit of [WWDL07], [LTL05], and [ZL11], and others. The key component of our sensor node model is a queueing system of M/M/1-type for the message queue management. Development of this queue is influenced by other processes which represent the specific features of the DSN. These processes will be considered as an environment for the queue with a **vice-versa** interaction.

We point out that we will make modeling simplifications as well, especially we will focus on the first key component, the data transmission, and will only include marginal parts of the queue management. To reduce technical effort we will discretize all state spaces. This simplification can be removed leading to Markov processes with general state spaces.

Related work: There are two main approaches for analytical modeling of DSN. The first is by direct construction of Markov processes and numerically solving the balance equation to obtain performance metrics from this. Typical examples are [CG06] and [CLWY07]. Jiang and Walrand studied the closely related CSMA protocol for the IEEE 802.11 Wireless Networks, and found explicit expressions for the stationary distribution of the network, see [JW10] and the references there. A single node and its environment is described in [Li11], exploiting matrix-geometrical methods.

The second approach is by utilizing stochastic network models with product form steady state. From this it is easy to obtain performance metrics. One often pays with oversimplification. But experience with the OR models for classical computer and communications systems is, that many systems are robust against deviations from, e.g., assumptions on service distributions. Using product form models usually one usually proceeds in the two-step construction, described above. For an idealized sensor network with sleep and active periods of the nodes Liu and Tong Lee [LTL05] applied this procedure. Mehmet Ali and Gu [MAG06] used a generalized Jackson network with unreliable nodes to model a DSN. From the product form steady state in [SD03b] the authors derive relevant performance metrics. Wu, Wang, Dang, and Lin [WWDL07] used a classical Jackson network to model "delay/fault-tolerant mobile sensor networks". A detailed analysis of a DSN with the aid of queueing network models is performed by Qiu, Feng, Xia, Wu, and Zhou [QFX$^+$11]. The networks are not of product form but similar to Mehmet Ali and Gu [MAG06] it is assumed that separability can be applied.

Our research in this paper is part of an ongoing project which focusses on investigations of queueing networks in a random environment. The aim is to find structures which show the asymptotic properties of separable networks (a) for the internal structure of Jackson-type or BCMP networks, and (b) for the interaction of the service network with the environment. Predecessors of our present work are e.g., [SD03a] (environment determines the availability of unreliable network nodes), [SSD$^+$06] (environment consists of an attached inventory, where stock-out lets the service process break down until replenishment arrives). A survey on related queueing-inventory systems is [KLM11]. Recent results on single nodes are in [KD12], [KD13b], with more relevant references.

The paper is structured as follows. In Section 2.2.5.2 we describe in detail the features of transmission and queue management protocols which we will incorporate into our model. Section 2.2.5.3 contains the main result on separability of the queue-environment interaction. In Section 2.2.5.4 we present details which can be further incorporated into the model, as well as possibilities to reduce complexity for easier computations. Some examples are presented in detail.

2.2.5.2. Model description

We consider a single mobile sensor node in a DSN (disruption tolerant wireless sensor network). Due to mobility and changing external conditions, this "referenced node" (RN) observes a varying environment with which the RN strongly interacts The functioning of the RN is governed by the following principles which select the relevant features and incorporate several interacting processes.

- Length of the packet queue of the RN ($\in \mathbb{N}_0$),
- number of active nodes in the one-hop neighborhood (the "outer environment") and the nodal delivery status of these neighbors,
- nodal delivery status of the RN (part of the "inner or local environment"),
- modus of the RN (active $= 1$, sleep $= 0$) (part of the "inner environment").

It follows that the referenced node RN can communicate with other nodes if and only if RN is **active** ($= 1$) **and** the number of active nodes in the outer environment is > 0 (for short we say, the outer environment then is **on** (> 0)).

When RN is active ($= 1$) and outer environment on (> 0), the stream of packets arriving at the packet queue of RN is the superposition of data gathering and receiving packets from other nodes. Following [WWDL07][p. 3291] we assume that the superposition process is a Poisson-λ process, and processing a packet in the queue needs an exponential-μ distributed time. The inter arrival times and service times are an independent set of variables.

Whenever RN is in sleep mode or the outer environment is off, all sensing, relaying, and sending activities of RN are frozen. This assumption is posed for simplicity, and is different from e.g. [LTL05], who allow during this periods data gathering by the node, but is in line with e.g. [ZL11].

Whenever RN is ready to send, it sends a packet to a one-hop neighbor. The routing decision is made on the basis of the nodal delivery status of the neighbors. The packet is send to the neighbor node with maximal nodal delivery value, say ζ, if there are ties these are broken by a pure random decision (with equal probability).

The nodal delivery values are on a scale $D := \{1, 2, \dots, d\}$ with $d < \infty$ the highest value. Whenever RN has send a packet to a node with delivery value ζ, it updates its own delivery value ξ as follows

$$\xi \to \begin{cases} \xi + 1, & \text{if } \xi < \zeta, \\ \xi - 1, & \text{if } \xi > \zeta, \\ \xi, & \text{if } \xi = \zeta. \end{cases} \tag{2.2.39}$$

Moreover, RN maintains a timer to adjust its nodal delivery value [WWDL07][p. 3288]: Whenever there has been no transmission within a time interval $\bar{\Delta}$, a timeout occurs and the delivery value ξ is updated (because RN could not transmit data during an interval $\bar{\Delta}$) as follows

$$\xi \to \begin{cases} \xi - 1, & \text{if } \xi > 1, \\ \xi, & \text{if } \xi = 1. \end{cases} \tag{2.2.40}$$

The length of the interval $\bar{\Delta}$ is Erlang-distributed with phase parameter δ and $T \geq 1$ phases (approximating a deterministic timer [WWDL07]). Whenever a transmission of RN occurs, the running timer is interrupted and immediately restarted. The phases are counted in decreasing order $T, T-1, \ldots, 1$. When the timer expires (the last phase, counted $= 1$, ends), it is reset to its maximal value T. The successive sampled timer intervals are independent and independent of the other activities.
Remark: In [WWDL07, p. 3288] the nodal delivery values are probabilities, i.e., on scale $[0,1]$, our rescaling is only for technical simplifications.

Message fault tolerances for the packets are introduced to estimate the importance of a stored (replication of a) message. Replications of some sent packets are stored for another transmission sometime later on. Whenever a copy of the sent message is stored at the end of the local queue, its message fault tolerance value is increased depending on the delivery values of the receiver. This will make the DSN less vulnerable against packet losses and will on the other hand not flood the network with too much messy packets.

Handling "message fault tolerance" is suppressed in the analytical single node model in [WWDL07]. We incorporate a simplified scheme in our model by the

Assumption on generating redundancy: Whenever RN has send a packet, it will store that packet with probability $f > 0$ at the end of its local queue. The storage decision is independent of the past. We assume in our first approach that $f = 0$ holds (as in [WWDL07]), $f > 0$ will be dealt with in Section 2.2.5.4.

Definition 2.2.20. The "outer environment process" describes the development of N nodes, which constitute the one-hop neighborhood of RN, and is assumed to be an irreducible homogeneous Markov process

$$O = (O(t) : t \geq 0), \qquad \text{with state space } \; E_o := (\{0\} \cup D)^N ,$$

where 0 in coordinate number k stands for "the k-th node is not available for RN", while $\eta_k > 0$ in coordinate k stands for "the k-th node is available for RN and has a nodal delivery value $\eta_k \in D$".

The generator of O is denoted by $Q_o = (q_o(y, y') : y, y' \in E_o)$ and the unique steady-state distribution of O is denoted by $\theta_o = (\theta_o(y) : y \in E_o)$.

We abbreviate for $\eta = (\eta_1, \ldots, \eta_N) \in E_o$: $g(\eta) := \max\{\eta_1, \ldots, \eta_N\}$, and shall say that the outer environment is "quiet" if $g = 0$ holds.

RN can communicate with other nodes iff RN is active $= 1$ **and** $g \neq 0$ holds.

Example 2.2.21. Several properties and components of the following environment process are taken from the model in [WWDL07]. (1) Reduction to a fixed cell where all other sensors are accessible, i.e., the one-hop neighborhood with N nodes is sufficient. (2) Independence of sensor nodes in the cell, which leads to (2.2.41) below and to the processes describing the behavior of the neighbors of RN are independent Markov processes with state space $\{0, 1, 2, \dots, d\}$. (3) Retrials are independent with identical success probability which is expressed in (2.2.41).

In the model [WWDL07, Section III] is not incorporated that as long as a node, say node j, is active its nodal delivery value evolves as a random walk on D as indicated by (2.2.39) and (2.2.40) which are taken from [WWDL07, Section II.A.(1)]. For simplicity, we assume that this random walk is Markov for its own, with upward jump rate $w_j^+(\eta_j)$ for node j in state $\eta_j < d$, and downward jump rate $w_j^-(\eta_j)$ in state $\eta_j > 1$. Its steady state is with normalization G_j

$$p_j(\eta_j) = G_j^{-1} \prod_{k=2}^{\eta_j} \frac{w_j^+(k-1)}{w_j^-(k)}, \quad \eta_j = 1, \dots, d.$$

With the help of these steady-state probabilities and constant rates $a_j, b_j > 0$ we incorporate additionally the active/sleep behavior (see Definition 2.2.22 below) of the nodes in the neighborhood into the positive local transition rates for node j as

$$q_{oj}(0, \eta_j) = b_j \cdot p_j(\eta_j), \quad \text{and} \quad q_{oj}(\eta_j, 0) = a_j, \quad \eta_j = 1, \dots, d\,.$$

The positive transition rates of O are for $j = 1, \dots, N$, and $\eta_j \in \{1, \dots, d\}$

$$\begin{aligned} &q_o((\eta_1, \dots, \eta_{j-1}, 0, \eta_{j+1}, \dots, \eta_N); (\eta_1, \dots, \eta_{j-1}, \eta_j, \eta_{j+1}, \dots, \eta_N)) = b_j \cdot p_j(\eta_j), \\ \text{and} \quad &q_o((\eta_1, \dots, \eta_{j-1}, \eta_j, \eta_{j+1}, \dots, \eta_N); (\eta_1, \dots, \eta_{j-1}, 0, \eta_{j+1}, \dots, \eta_N)) = a_j. \end{aligned} \tag{2.2.41}$$

The environment process O is ergodic with steady-state probabilities

$$\theta_o(\eta_1, \dots, \eta_N) = \prod_{j=1}^{N} \left(\frac{a_j}{a_j + b_j}\right)^{1_{(\eta_j = 0)}} \left(\frac{b_j}{a_j + b_j} \cdot p_j(\eta_j)\right)^{1_{(\eta_j \neq 0)}}, (\eta_1, \dots, \eta_N) \in E_o.$$

The outer environment is quiet with probability $\theta_o(0, \dots, 0) = \prod_{j=1}^{N}(a_j/(a_j + b_j))$, and for $k = 1, \dots, d$ the probability that $\{g \geq k\}$ holds is

$$1 - \prod_{j=1}^{N} \left(\frac{a_j}{a_j + b_j} + \frac{b_j}{a_j + b_j} \sum_{\eta_j = 1}^{k-1} p_j(\eta_j)\right). \tag{2.2.42}$$

Definition 2.2.22. The inner (local) environment of RN is a stochastic process, which is not Markov for its own

$$I = (I(t) : t \geq 0), \qquad \text{with state space} \quad E_i := \Delta \times D,$$

where $(t, \xi) \in \Delta \times D$ indicates that the timer is in phase t and the nodal delivery status of RN is ξ. Recall: 1 stands for lowest, d for highest delivery value, $\Delta := \{1, \dots, T\}$ are the possible residual exponential-δ phases of the running timer.

Active and sleep phases of RN are governed by an alternating renewal process

$$A = (A(t) : t \geq 0), \qquad \text{with state space} \quad E_a := \{0, 1\},$$

where 1 stands for "active" and 0 stands for "sleep". The dwell time in the active status is exponential-α, whereas in sleep status exponential-β. The unique steady-state distribution of A is $\theta_a = (\theta_a(0), \theta_a(1)) = (\beta/(\alpha + \beta), \alpha/(\alpha + \beta))$.

During RN's sleep times all its activities are frozen: Sending, receiving, timer.

Definition 2.2.23. The queue length process of RN is a process, which is not Markov for its own

$$X = (X(t) : t \geq 0), \qquad \text{with state space } \mathbb{N}_0,$$

where $X(t)$ counts the number of packets stored in RN, either under transmission (in service) or waiting. Whenever RN is sleeping or the outer environment is quiet, there is no service possible,the packet on the service place is stored there, and no new arrival is admitted until RN becomes active again **and** there are active nodes in the outer environment.

Note, that whenever RN is in active mode, its timer is running, irrespective of the status of the outer environment O.

Assumption 2.2.24. *We make the following natural assumption: The processes O and A are independent and independent of the set of inter arrival and service times, and of the timer intervals.*

Note, that independence of A and O from the timer is for free if the timer is deterministic. A direct consequence of the definitions and Assumption 2.2.24 is

Proposition 2.2.25. *With $Y := (A, I, O)$ the process $Z := (X, Y) = (X, A, I, O)$ is a homogeneous Markov process, which is irreducible on state space*

$$E := \mathbb{N}_0 \times E_a \times E_i \times E_o = \mathbb{N}_0 \times \{0, 1\} \times \Delta \times D \times (\{0\} \cup D)^N.$$

We denote the generator of Z by $Q = (q(z, z') : z, z' \in E)$, and, when it exists, the (then uniquely defined) steady-state distribution of Z, by $\pi = (\pi(z) : z \in E)$.

In general the queue length process X and its environment Y are strongly dependent. In one direction, the environment can shut down the service and arrival process, while in the other direction transmission of a message changes the nodal delivery value and resets the timer. We emphasize that in the first case the environment changes and the queue length stays at its present value, while in the second case the queue length **and** the environment jump concurrently. This property is discussed in Remark 2.2.28 below in comparison with the literature.

2.2.5.3. Steady-state behavior

We assume in the following that the node and its environment can stabilize in the long run, i.e., the joint process $Z = (X, Y)$ is ergodic, which implies that a steady state of Z exist. Recall from p. 68: When RN is able to communicate, the arrival rate is λ, the service rate is μ.

Proposition 2.2.26. *Z is ergodic on E iff $\lambda < \mu$. Then its unique stationary distribution π fulfills for all $(n, a, t, \xi, \eta) \in \mathbb{N}_0 \times \{0, 1\} \times \Delta \times D \times (\{0\} \cup D)^N$*

$$\pi(n, a, t, \xi, \eta) = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^n \cdot \theta(a, t, \xi, \eta), \tag{2.2.43}$$

where θ is a (unique) probability distribution on $\{0, 1\} \times \Delta \times D \times (\{0\} \cup D)^N$.

Proof. In the balance equations for Z we abbreviate $-q_o(\eta,\eta) =: q_o(\eta), \eta \in E_o$, and when no restriction is posed, a variable runs through all admissible values. Recall $g(\eta) := \max(\eta_1, \dots, \eta_N)$.
For $g(\eta) \neq 0$, $t < T$, $a = 1$:

$$\begin{aligned} &\pi(n,1,t,\xi,\eta)[\lambda + \mu 1_{(n>0)} + \alpha + \delta + q_o(\eta)] \\ &= \pi(n-1,1,t,\xi,\eta)\lambda 1_{(n>0)} + \pi(n,0,t,\xi,\eta)\beta + \pi(n,1,t+1,\xi,\eta)\delta \\ &\quad + \sum_{\gamma \in E_o \setminus \{\eta\}} \pi(n,1,t,\xi,\gamma) q_o(\gamma,\eta). \end{aligned} \tag{2.2.44}$$

For $g(\eta) \neq 0$, $a = 1$:

$$\begin{aligned} &\pi(n,1,T,\xi,\eta)[\lambda + \mu 1_{(n>0)} + \alpha + \delta + q_o(\eta)] \\ &= \pi(n-1,1,T,\xi,\eta)\lambda 1_{[n>0]} + \pi(n,0,T,\xi,\eta)\beta + \sum_{\gamma \in E_o \setminus \{\eta\}} \pi(n,1,T,\xi,\gamma) q_o(\gamma,\eta) \\ &\quad + \pi(n,1,1,\xi,\eta)\delta 1_{[\xi=1]} + \pi(n,1,1,\xi+1,\eta)\delta 1_{[\xi<d]} + \sum_{s=1}^{T} \pi(n+1,1,s,\xi,\eta)\mu 1_{[\xi=g(\eta)]} \\ &\quad + \sum_{s=1}^{T} \pi(n+1,1,s,\xi-1,\eta)\mu 1_{[0<\xi-1<g(\eta)]} \\ &\quad + \sum_{s=1}^{T} \pi(n+1,1,s,\xi+1,\eta)\mu 1_{[d\geq\xi+1>g(\eta)>0]}. \end{aligned} \tag{2.2.45}$$

For η with $g(\eta) = 0$, $t < T$, $a = 1$:

$$\begin{aligned} &\pi(n,1,t,\xi,\eta)[\alpha + \delta + q_o(\eta)] \\ &= \pi(n,0,t,\xi,\eta)\beta + \pi(n,1,t+1,\xi,\eta)\delta + \sum_{\gamma \in E_o \setminus \{\eta\}} \pi(n,1,t,\xi,\gamma) q_o(\gamma,\eta). \end{aligned} \tag{2.2.46}$$

For η with $g(\eta) = 0$, $a = 1$:

$$\begin{aligned} \pi(n,1,T,\xi,\eta)[\alpha + \delta + q_o(\eta)] &= \pi(n,0,T,\xi,\eta)\beta + \sum_{\gamma \in E_o \setminus \{\eta\}} \pi(n,1,T,\xi,\gamma) q_o(\gamma,\eta) \\ &\quad + \pi(n,1,1,\xi,\eta)\delta 1_{[\xi=1]} + \pi(n,1,1,\xi+1,\eta)\delta 1_{[\xi<d]}. \end{aligned} \tag{2.2.47}$$

For $a = 0$:

$$\pi(n,0,t,\xi,\eta)[\beta + q_o(\eta)] = \pi(n,1,t,\xi,\eta)\alpha + \sum_{\gamma \in E_o \setminus \{\eta\}} \pi(n,0,t,\xi,\gamma) q_o(\gamma,\eta). \tag{2.2.48}$$

Inserting $\pi(n,a,t,\xi,\eta) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \cdot \theta(a,t,\xi,\eta)$ into equations (2.2.44)–(2.2.48) reveals that the queue length terms $\left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n$ cancel completely, which yields the following set of reduced equations.
For $g(\eta) \neq 0$, $t < T$, $a = 1$:

$$\theta(1,t,\xi,\eta)[\lambda + \alpha + \delta + q_o(\eta)] = \theta(0,t,\xi,\eta)\beta + \theta(1,t+1,\xi,\eta)\delta + \sum_{\gamma \in E_o \setminus \{\eta\}} \theta(1,t,\xi,\gamma) q_o(\gamma,\eta).$$

For $g(\eta) \neq 0$, $a = 1$:

$$\begin{aligned}
&\theta(1,T,\xi,\eta)[\lambda + \alpha + \delta + q_o(\eta)] \\
&= \theta(0,T,\xi,\eta)\beta + \sum_{\gamma \in E_o \setminus \{\eta\}} \theta(1,T,\xi,\gamma)q_o(\gamma,\eta) + \theta(1,1,\xi,\eta)\delta 1_{[\xi=1]} \\
&\quad +\theta(1,1,\xi+1,\eta)\delta 1_{[\xi<d]} + \sum_{s=1}^{T} \theta(1,s,\xi,\eta)\lambda 1_{[\xi=g(\eta)]} \\
&\quad + \sum_{s=1}^{T} \theta(1,s,\xi-1,\eta)\lambda 1_{[0<\xi-1<g(\eta)]} + \sum_{s=1}^{T} \theta(1,s,\xi+1,\eta)\lambda 1_{[d\geq\xi+1>g(\eta)>0]}.
\end{aligned}$$

For η with $g(\eta) = 0$, $t < T$, $a = 1$:

$$\theta(1,t,\xi,\eta)[\alpha + \delta + q_o(\eta)] = \theta(0,t,\xi,\eta)\beta + \theta(1,t+1,\xi,\eta)\delta + \sum_{\gamma \in E_o \setminus \{\eta\}} \theta(1,t,\xi,\gamma)q_o(\gamma,\eta).$$

For η with $g(\eta) = 0$, $a = 1$:

$$\begin{aligned}
\theta(1,T,\xi,\eta)[\alpha + \delta + q_o(\eta)] = \theta(0,T,\xi,\eta)\beta &+ \sum_{\gamma \in E_o \setminus \{\eta\}} \theta(1,T,\xi,\gamma)q_o(\gamma,\eta) \\
&+\theta(1,1,\xi,\eta)\delta 1_{[\xi=1]} + \theta(1,1,\xi+1,\eta)\delta 1_{[\xi<d]}.
\end{aligned}$$

For $a = 0$:

$$\theta(0,t,\xi,\eta)[\beta + q_o(\eta)] = \theta(1,t,\xi,\eta)\alpha + \sum_{\gamma \in E_o \setminus \{\eta\}} \theta(0,t,\xi,\gamma)q_o(\gamma,\eta).$$

With elementary, but tedious computations it can be shown that this is a "generator equation", i.e., there exists some continuous time Markov process on the finite state space $K := \{0,1\} \times \Delta \times D \times (\{0\} \cup D)^N$ with generator matrix $Q_{\text{red}} = (q_{\text{red}}(y,y') : y, y' \in K$ such that the reduced system of equations is $\theta \cdot Q_{\text{red}} = 0$. The main effort is to show that the row sums of Q_{red} are zero.

This generator equation has a unique probability solution because K is finite and Q_{red} is irreducible. $\square$

The result of Proposition 2.2.26 is surprising. Obviously, the queueing process X which is the central unit of the message handling and transmission management system and the environment process Y strongly interact. Nevertheless, in steady state and in the long run the joint steady-state distribution for a fixed time instant is the independent coupling of the respective marginal steady-state distributions. This resembles the independence of the marginal queue lengths in a stationary Jackson network [Jac57]. The new feature here is that X and Y are processes of very different structure, while in Jackson's theorem the queue lengths are processes of similar nature.

Similarly, as it is well known in the case of Jackson networks, our result does not say that X and Y are independent processes. There are correlations over time in $(Z(s), Z(t))$ for $0 \leq s < t$ and for different time instants there are correlations between $X(s)$ and $Y(t)$. The investigation of this correlation structure is part of our ongoing research.

Another remarkable property of the system following from Proposition 2.2.26 is an invariance property: Whenever for a pair λ, μ with $\lambda < \mu$ we have computed the marginal environment steady state θ, this is the θ as function of λ (not of μ!) for all other pairs with $\lambda < \mu$. This is of interest in cases of a complicated environment, where $\theta \cdot Q_{\text{red}} = 0$ may be not easy to obtain.

Remark 2.2.27. From the very definition of the active-sleep process A and the outer environment process O and Assumption 2.2.24 it follows that the marginal distribution of A is θ_a given in Definition 2.2.22 and the marginal distribution of O is θ_o indicated in Definition 2.2.20.

The solution θ of $\theta \cdot Q_{\text{red}} = 0$, found in the proof of Proposition 2.2.26 in general does not factorize further. But even if we cannot factorize θ further it is helpful by reducing an infinite linear system of equations to a finite system.

Remark 2.2.28. Boucherie [Bou94] considered vector processes of independent coordinates, where restriction on the transitions are imposed as follows: A coordinate process, say S_j, by entering a specified subset A_j of its state space, where he competes with a second process, say S_k for resources (which can be used by only one process at a time) shut down S_k completely, as long as it stays in A_j. This is similar to our "vector process". The difference is: In [Bou94] it is assumed that only one coordinate of the vector process can change at time, and the starting point are independent Markov processes. Neither property is required here: Not all processes used for the construction are independent, and there occur simultaneous jumps of the queue and the environment, as can be seen in (2.2.45).

2.2.5.4. Extensions and refinements

Modeling the outer environment

The Markov process O to describe the development of the outer environment is constructed in Definition 2.2.20 in the spirit of the neighborhood construction of [WWDL07]. Note, that there dynamics of the nodal delivery values are substituted by a fixed value. Into our process O we have incorporated dynamics of nodal delivery values without much effort, still obtaining explicit expressions.

Our modeling procedure offers to incorporate much more versatile dynamical schemes. This can be seen from the proof of Proposition 2.2.26: It is not necessary that the set $\{g = 0\} \subset E_o$ is single valued. There may be more states of the outer environment which do not allow RN to communicate with its neighbors for different reasons.

On the other side this flexibility offers model reductions. Starting from a complex environment space E_o' with rates $q_o'(k,\ell)$ and some "decision function" (other than the simple maximum) $g' : U' \to \{0, 1, \ldots, d\}$ we can reduce complexity via $U'' := g(U') := \{0, 1, \ldots, d\}$ and assume that the functional process $g(O)$ is Markovian itself. Reasonable (approximate) transition rates then are

$$q_o''(k,\ell) := S(k,\ell)^{-1} \sum_{\eta \in U, g'(\eta)=k} \sum_{\zeta \in U, g'(\zeta)=\ell} q_o'(\eta,\zeta),$$

where $S(k,\ell) = \sum_{\eta \in U', g'(\eta)=k} \sum_{\zeta \in U', g'(\zeta)=\ell} 1_{[q_o'(\eta,\zeta)>0]}$ is the number of positive transitions from $\{g' = k\}$ to $\{g' = \ell\}$.

More reduction is obtained if we distinguish in the status of the outer environment only states 0 (= no active neighbor = quiet outer environment) and 1 (= at least one active neighbor).

In any case: The proof of Proposition 2.2.26 applies without changes.

Example 2.2.29. Consider the outer environment E_o from Definition 2.2.20. Then $g(E_o) = \{0, 1, \dots, d\}$ and from (2.2.42) the probability of $\{g = 0\}$ is $p(0) = \prod_{j=1}^{N}(a_j/(a_j + b_j))$, while the maximal nodal delivery value in the neighborhood is $k \geq 1$ (i.e., $\{g = k\}$) with probability $p(k) :=$

$$\prod_{j=1}^{N}\left(\frac{a_j}{a_j+b_j} + \frac{b_j}{a_j+b_j}\sum_{\eta_j=1}^{k} p_j(\eta_j)\right) - \prod_{j=1}^{N}\left(\frac{a_j}{a_j+b_j} + \frac{b_j}{a_j+b_j}\sum_{\eta_j=1}^{k-1} p_j(\eta_j)\right).$$

A kernel to generate dynamics for this equilibrium is with positive transition rates

$$q(i, i+1) = \ell(i), \quad i = 0, 1, \dots, d-1, \quad \text{and} \quad q(i, i-1) = m(i), \quad i = 1, \dots, d,$$

with $\ell(i) = 1/p(i)$ for $i = 0, 1, \dots, d-1$ and $m(i) = 1/p(i)$ for $i = 1, \dots, d$. This yields a reversible dynamics with the required target distribution.

Example 2.2.30. Consider the outer environment E_o from Definition 2.2.20 with $D := \{1\}$, i.e., we distinguish only whether the N nodes of RN's one-hop neighborhood are available or not. Then $E_o = \{(0, 1)\}^N$ and the positive transition rates for O are, for $j = 1, \dots, N$, and $\eta_j = 1$

$$\begin{aligned} &q_o((\eta_1, \dots, \eta_{j-1}, 0, \eta_{j+1}, \dots, \eta_N); (\eta_1, \dots, \eta_{j-1}, \eta_j, \eta_{j+1}, \dots, \eta_N)) = b_j \\ \text{and} \quad &q_o((\eta_1, \dots, \eta_{j-1}, \eta_j, \eta_{j+1}, \dots, \eta_N); (\eta_1, \dots, \eta_{j-1}, 0, \eta_{j+1}, \dots, \eta_N)) = a_j. \end{aligned}$$

The environment process O is ergodic and the steady-state probabilities are

$$\theta_o(\eta_1, \dots, \eta_N) = \prod_{j=1}^{N}\left(\frac{a_j}{a_j+b_j}\right)^{1_{[\eta_j=0]}}\left(\frac{b_j}{a_j+b_j}\right)^{1_{[\eta_j=1]}}, \quad \forall(\eta_1, \dots, \eta_N) \in E_o.$$

The outer environment is quiet with probability $\theta_o(0, \dots, 0) = \prod_{j=1}^{N}(a_j/(a_j + b_j))$.

Modeling fault tolerance

The introduction of fault tolerance values for any message and its updating in course of transmitting a message and possibly restoring it in the message queue of RN tries to support the resilience of the network without flooding it with messages. Modeling this in a detailed way would need to introduce for the messages different types which change over time and, if we follow the details of the protocol in [WWDL07], type-dependent priorities and reordering of the packets according to the fault tolerance values. As the authors in that paper noticed, such scheme probably can not be modeled analytically in full detail. So message replication is skipped in their model.

A simple way to incorporate the effect of increasing queue lengths by a randomized message replication is to estimate an overall replication probability $f \in [0, 1]$ for sent messages and consider the message queue as a feedback queue: If a message is served, it is fed back with feedback probability $f > 0$ to the tail of the queue. We immediately obtain the

Corollary 2.2.31. *Let the message queue be a feedback queue with feedback probability $f > 0$. Z is ergodic on E iff $\lambda < \mu(1-f)$. Then its unique stationary distribution π_f fulfills for all $(n, a, t, \xi, \eta) \in \mathbb{N}_0 \times \{0,1\} \times \Delta \times D \times (\{0\} \cup D)^N$*

$$\pi_f(n, a, t, \xi, \eta) = \left(1 - \frac{\lambda}{\mu(1-f)}\right)\left(\frac{\lambda}{\mu(1-f)}\right)^n \cdot \theta(a, t, \xi, \eta),$$

with θ the distribution on $\{0,1\} \times \Delta \times D \times (\{0\} \cup D)^N$ from Proposition 2.2.26.

The result says that with replication probability f the load of the referenced node is the same as without replication but with a prolongation of the transmission time according to service rate $\mu(1-f)$.

Although with the corollary we are in a position to adapt the load of the message queue better to the situation with message replication we have to pay for this with a slight drawback, which may be not obvious. The protocol in [WWDL07] declares that the nodal delivery value is updated every time a message is sent out. In the model described in Corollary 2.2.31 updating is formally done only when a message is sent out and no feedback occurs.

Without going into the details we mention only that we can remedy this drawback by introducing an additional update process which updates RN's nodal delivery value according to the scheme (2.2.39) at time points generated by a Poisson-μf process when RN is active.

Intensity of the arrival process

In [WWDL07], [LTL05], [ZL11]), the arrivals at the message queue of RN are assumed to be generated by two independent Poisson processes. A Poisson-r process generates the data for RN, while a Poisson-ℓ process is generated by the nodes in the neighborhood of RN. The intensity of the Poisson-ℓ process is determined by the states of the outer environment. ℓ is typically computed as a gross value on the basis of the environments steady state. An example is given in [WWDL07, p. 3291]. In a similar way we can, starting from the information decoded in the distribution θ from Proposition 2.2.26, estimate the overall arrival rate ℓ at RN from the outside. We then set $\lambda := r + \ell$.

Reducing the dimension of the environment process

The reduction of complexity described above leave the dimension of the state space of Z invariant while diminishing the sizes of components. Further reduction can be obtained by incorporating the effects of the outer environment into the transition regime of the inner environment and by canceling thereafter the component process O. The resulting process $Z := (X, Y) = (X, A, I)$ will be a homogeneous irreducible strong Markov process on $E := \mathbb{N}_0 \times E_a \times E_i = \mathbb{N}_0 \times \{0,1\} \times \Delta \times D$. The development of the nodal delivery status of RN is governed by the timer $\bar{\Delta}$ as before via (2.2.40) and by the rule that whenever RN has send a packet, it updates its delivery value ξ as follows

$$\xi \to \begin{cases} \xi + 1, & \text{with probability } r^+(\xi) \text{ if } \xi < d, \\ \xi - 1, & \text{with probability } r^-(\xi) \text{ if } \xi > 1, \\ \xi, & \text{with probability } 1 - r^+(\xi)1_{[\xi<d]} - r^-(\xi)1_{[\xi>1]} \text{ if } 1 \le \xi \le d. \end{cases}$$

As in Proposition 2.2.26 then follows for the system with no replication of sent messages ($f = 0$) the first part of the next statement, while the second part is again surprising.

Corollary 2.2.32. *Z is ergodic on the reduced state space $E = \mathbb{N}_0 \times E_a \times E_i = \mathbb{N}_0 \times \{0,1\} \times \Delta \times D$ iff $\lambda < \mu$. Then its unique stationary distribution π fulfills*

$$\pi(n,a,t,\xi) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \cdot \theta(a,t,\xi), \quad \forall (n,a,t,\xi) \in \mathbb{N}_0 \times \{0,1\} \times \Delta \times D,$$

where θ is a uniquely defined probability distribution on $\{0,1\} \times \Delta \times D$.

Moreover, θ factorizes completely for $(a,t,\xi) \in \{0,1\} \times \Delta \times D$ according to

$$\theta(a,t,\xi) = \left(\frac{\beta}{\alpha+\beta}\right)^a \cdot \left(\frac{\alpha}{\alpha+\beta}\right)^{1-a} \cdot \left(\frac{\delta}{\lambda+\delta}\right)^{T-t} K_\Delta^{-1} \cdot \psi(\xi). \tag{2.2.49}$$

Here $K_\Delta = \frac{\delta}{\lambda}\left(\frac{\lambda+\delta}{\delta} - \left(\frac{\delta}{\lambda+\delta}\right)^{T-1}\right)$ is the normalization for the timer distribution, and ψ is a probability on D, the marginal nodal delivery value distribution.

Proof. Whenever there is no restriction for a variable indicated, it runs through all admissible values. The steady-state equations for Z are then
For $t < T$, $a = 1$:

$$\begin{aligned} &+\pi(n,1,t,\xi)[\lambda + \mu 1_{[n>0]} + \alpha + \delta] \\ &= \pi(n-1,1,t,\xi)\lambda 1_{[n>0]} + \pi(n,0,t,\xi)\beta + \pi(n,1,t+1,\xi)\delta. \end{aligned}$$

For $a = 1$:

$$\begin{aligned} &\pi(n,1,T,\xi)[\lambda + \mu 1_{[n>0]} + \alpha + \delta] \\ &= \pi(n-1,1,T,\xi)\lambda 1_{[n>0]} + \pi(n,1,1,\xi)\delta 1_{[\xi=1]} \\ &+ \sum_{s=1}^{T} \pi(n+1,1,s,\xi-1)\mu r^+(\xi-1)1_{[\xi>1]} + \pi(n,1,1,\xi+1)\delta 1_{[\xi<d]} \\ &+ \sum_{s=1}^{T} \pi(n+1,1,s,\xi+1)\mu r^-(\xi+1)1_{[\xi<d]}\pi(n,0,T,\xi)\beta \\ &+ \sum_{s=1}^{T} \pi(n+1,1,s,\xi)\mu\left[1 - r^+(\xi)1_{[\xi<d]} - r^-(\xi)1_{[\xi>])}\right]. \end{aligned}$$

For $a = 0$:

$$\pi(n,0,t,\xi)\beta = \pi(n,1,t,\xi)\alpha.$$

Inserting $\pi(n,a,t,\xi) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \cdot \left(\frac{\beta}{\alpha+\beta}\right)^a \cdot \left(\frac{\alpha}{\alpha+\beta}\right)^{1-a} \cdot \phi(t,\xi)$, where ϕ is a function of (t,ξ) only, into these equations reveals that the terms $\left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \cdot \left(\frac{\beta}{\alpha+\beta}\right)^a \cdot \left(\frac{\alpha}{\alpha+\beta}\right)^{1-a}$ cancel completely, yielding a set of reduced equations:
For $t < T$:

$$\phi(t,\xi)[\lambda + \delta] = \phi(t+1,\xi)\delta,$$

$$\begin{aligned}\phi(T,\xi)[\lambda+\delta] = \phi(1,\xi)\delta 1_{[\xi=1]} &+ \sum_{s=1}^{T}\phi(s,\xi-1)\lambda r^{+}(\xi-1)1_{[\xi>1]} \\ &+\phi(1,\xi+1)\delta 1_{[\xi<d]} + \sum_{s=1}^{T}\phi(s,\xi+1)\lambda r^{-}(\xi+1)1_{[\xi<d]} \\ &+\sum_{s=1}^{T}\phi(s,\xi)\lambda\left[1-r^{+}(\xi)1_{[\xi<d]}-r^{-}(\xi)1_{[\xi>1]}\right].\end{aligned}$$

For $t=1,\ldots,T$, the first equation yields $\phi(t,\xi)=\phi(T,\xi)\left(\delta/(\lambda+\delta)\right)^{T-t}$, and we set $\phi(t,\xi)=\left(\frac{\delta}{\lambda+\delta}\right)^{T-t}K_{\Delta}^{-1}\cdot\psi(\xi)$ for some function $\psi(\xi)$. The first equation is solved obviously by this expression, and the second turns into $\psi(\xi)[\lambda+\delta]=$

$$\begin{aligned}&\psi(\xi)\left(\frac{\delta}{\lambda+\delta}\right)^{T-1}\delta 1_{[\xi=1]} + \sum_{s=1}^{T}\psi(\xi-1)\left(\frac{\delta}{\lambda+\delta}\right)^{T-s}\lambda r^{+}(\xi-1)1_{[\xi>1]} \\ &+\psi(\xi+1)\left(\frac{\delta}{\lambda+\delta}\right)^{T-1}\delta 1_{[\xi<d]} + \sum_{s=1}^{T}\psi(\xi+1)\left(\frac{\delta}{\lambda+\delta}\right)^{T-s}\lambda r^{-}(\xi+1)1_{[\xi<d]} \\ &+\sum_{s=1}^{T}\psi(\xi)\left(\frac{\delta}{\lambda+\delta}\right)^{T-s}\lambda\left[1-r^{+}(\xi)1_{[\xi<d]}-r^{-}(\xi)1_{[\xi>1]}\right].\end{aligned}$$

Recall that $\sum_{s=1}^{T}\left(\frac{\delta}{\lambda+\delta}\right)^{T-s}=K_{\Delta}$ and $\frac{1}{\lambda+\delta}K_{\Delta}\lambda=\delta\left[\frac{\lambda+\delta}{\delta}-\left(\frac{\delta}{\lambda+\delta}\right)^{T-1}\right]$. Dividing by $\lambda+\delta$ and utilizing this property yields $\psi(\xi)=$

$$\begin{aligned}&\psi(\xi)\left(\frac{\delta}{\lambda+\delta}\right)^{T}1_{[\xi=1]} + \psi(\xi-1)\left(1-\left(\frac{\delta}{\lambda+\delta}\right)^{T}\right)r^{+}(\xi-1)1_{[\xi>1]} \\ &+\psi(\xi+1)\left(\frac{\delta}{\lambda+\delta}\right)^{T}1_{[\xi<d]} + \psi(\xi+1)\left(1-\left(\frac{\delta}{\lambda+\delta}\right)^{T}\right)r^{-}(\xi+1)1_{[\xi<d]} \\ &+\psi(\xi)\left(1-\left(\frac{\delta}{\lambda+\delta}\right)^{T}\right)\left[1-r^{+}(\xi)1_{[\xi<d]}-r^{-}(\xi)1_{[\xi>1]}\right].\end{aligned}$$

Taking $\psi(1)$ as unknown, this is a two-term recursion for the $\psi(\xi)$, which are uniquely determined up to the factor $\psi(1)$, which is determined from $\psi(1)+\ldots+\psi(d)=1$. This must hold, because the proposed product form for θ, respectively π, is by ergodicity of Z a probability. □

Comments: The marginal timer distribution reveals that the most probable timer value is T. The geometrical decay of the residual timer state probabilities is faster when the arrival intensity increases. This reflects the timer policy: The timer is reset to T whenever a message is sent out.

It is not intuitive that (i) the timer distribution is independent of α and β, because the timer is interrupted whenever RN is in sleep mode, and (ii) the timer and the active-sleep processes are at fixed times independent.

Corollary 2.2.33. *In the setting of Corollary 2.2.32 in steady state the throughput of RN is*

$$TH(RN) = \lambda \left(1 - \frac{\beta}{\alpha+\beta} - \psi(0)\frac{\alpha}{\alpha+\beta}\right).$$

2.2.6. Crusher station in open-pit mining

In this section we will analyze a crusher station from [ZW09, Section 2] and approximate it by a loss system. This crusher station is a part of an open pit mining system, were loaded trucks drive to two discharging positions. There, the trucks discharge their loads into a bin with a crusher and drive away. The bin buffers the loads between trucks and the crusher. The capacity of the bin is limited to N loads. The capacity of the crusher is limited to a single truck load. When the bin and the crusher are full, the discharging stations are blocked. When discharging position is occupied by a truck or is blocked, new arriving trucks wait on the road forming a queue. Each truck carries a single load, therefore for the whole system it is sufficient to focus only on a flow of truck loads from arriving into the queue at discharging positions until complete processing by the crusher.

The original queuing model from [ZW09, Section 2] is a tandem system where a first node of type $M/M/2/\infty$ models two discharging positions and the second node of type $M/D/1/N$ models[14] a crusher with capacity 1 and a bin with capacity N. See Figure 2.2.8 on page 80. The authors discovered by stochastic simulations that for a system with $N = 2$

> "... with the defined parameters, the mean number of trucks in the queue starts increasing sharply after the average truck arrival rate reaches 28 per hour, and approaches a very high number when the average truck arrival rate reaches the throughput of the crusher, 35 truck loads per hour." [ZW09, p. 3]

The aim of this section is to reproduce this qualitative behavior using a loss-system approximation from Section 2.1.4.

The original model in [ZW09, Section 2] has two discharging positions, each with service rate μ_{orgn} and exponential service times, and a crusher with service rate ν and deterministic service times. Instead of this, we will use a simplified version with single discharging position with service rate $\mu = 2 \cdot \mu_{\text{orgn}}$ and a crusher with exponential service times. This simplification allows us to use loss-system approximation developed in Section 2.1.4.

2.2.6.1. Model without loss

Recall that we use subscript NL (**N**o **L**oss) for the systems whose costumers are not lost due to blocking.

Definition 2.2.34 (Crusher station without loss)**.** Given a tandem system with two nodes. The service times at both nodes are exponential. The stochastic process

[14] In [ZW09] the $\cdot/\cdot/1/N$ queues have N maximal number of customers. In contrast to [ZW09], here, $\cdot/\cdot/1/N$ queues have N waiting positions and their maximal number of customers is $N + 1$.

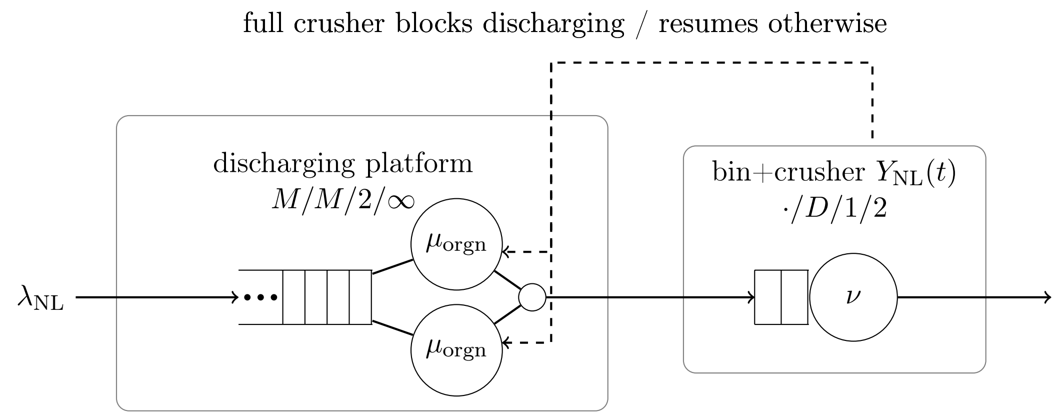


Figure 2.2.8.: Original model of a crusher station in open-pit mining with maximal capacity of the second node $N = 2$.

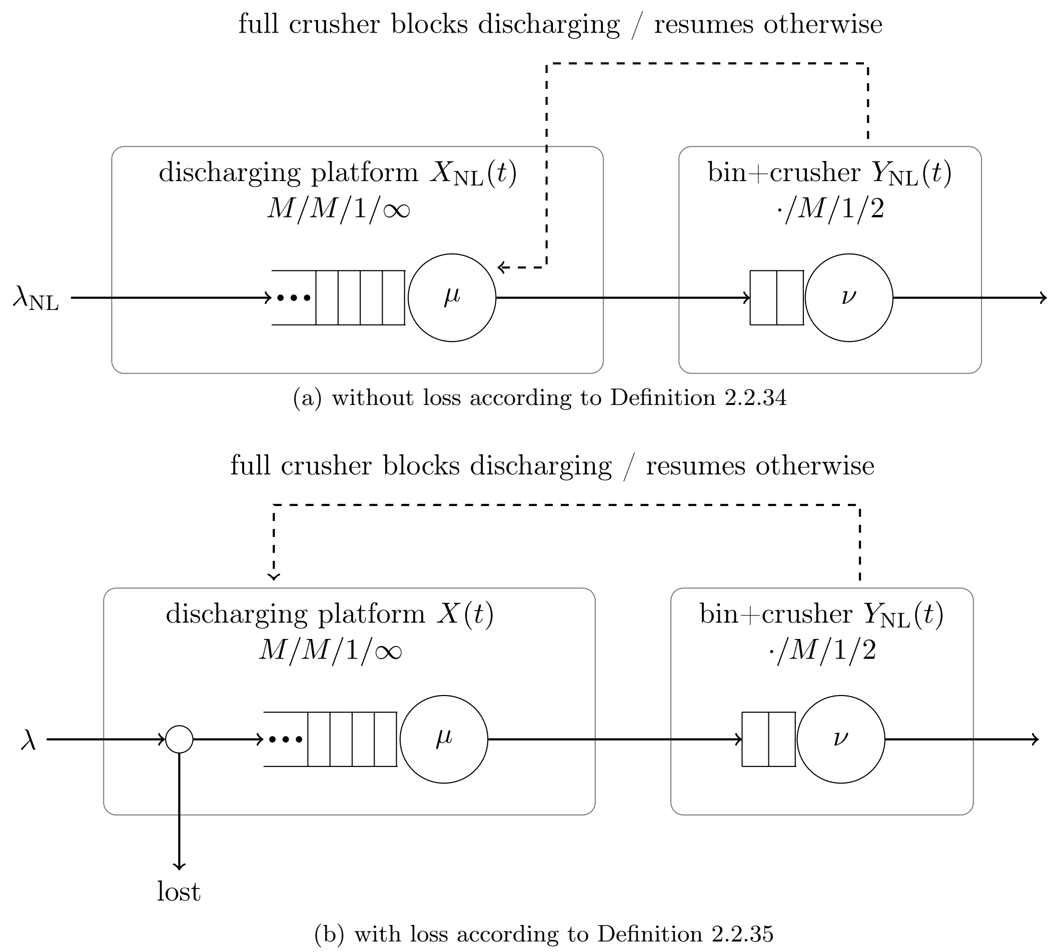


Figure 2.2.9.: Model of a crusher in station.

$(X_{\mathrm{NL}}(t) : t \geq 0)$ with $X_{\mathrm{NL}}(t) \in \mathbb{N}_0$ describes the number of truck loads at the first node. The stochastic process $(Y_{\mathrm{NL}}(t) : t \geq 0)$ with $Y_{\mathrm{NL}}(t) \in K := \{0, 1, 2, 3\}$ describes number of truck loads in the second node. The first node models a discharging platform with service rate μ and infinite number of waiting places. The second node models a crusher with service rate ν with two additional waiting places in a bin. It can contain maximal three truck loads. The truck loads arrive in a Poisson stream with rate λ_{NL} at the first node. When the second node is full, the first node interrupts service. This blocking mechanisms is called *blocking-before-service* [Per90, p. 455]. See Figure 2.2.9a on page 80. When the first node is blocked, newly arriving truck loads are *not* lost, but wait in the queue of the first node.

The whole system is described by a continues-time Markov process $Z_{\mathrm{NL}} = ((X_{\mathrm{NL}}(t), Y_{\mathrm{NL}}(t))) : t \geq 0)$ with generator $Q_{\mathrm{NL}} = (q(i,j) : i,j \in \mathbb{N} \times K) =$

$$
\left(\begin{array}{c|cccc|cccc|cccc|c}
 & (0,0) & (0,1) & (0,2) & (0,3) & (1,0) & (1,1) & (1,2) & (1,3) & (2,0) & (2,1) & (2,2) & (2,3) & \dots \\
\hline
(0,0) & \star & & & & \lambda_{\mathrm{NL}} & & & & & & & & \\
(0,1) & \nu & \star & & & & \lambda_{\mathrm{NL}} & & & & & & & \\
(0,2) & & \nu & \star & & & & \lambda_{\mathrm{NL}} & & & & & & \\
(0,3) & & & \nu & \star & & & & \lambda_{\mathrm{NL}} & & & & & \\
\hline
(1,0) & & \mu & & & \star & & & & \lambda_{\mathrm{NL}} & & & & \\
(1,1) & & & \mu & & \nu & \star & & & & \lambda_{\mathrm{NL}} & & & \\
(1,2) & & & & \mu & & \nu & \star & & & & \lambda_{\mathrm{NL}} & & \\
(1,3) & & & & & & & \nu & \star & & & & \lambda_{\mathrm{NL}} & \\
\hline
(2,0) & & & & & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & \dots \\
\vdots & & & & & & & & & & & & &
\end{array}\right).
$$

Here $\star$ is a placeholder for diagonal elements calculated according to the formula $q_{\mathrm{NL}}(i,i) = -\sum_{j \in (\mathbb{N}_0 \times K) \setminus \{i\}} q(i,j)$.

The generator Q_{NL} has a block-tridiagonal form

$$
Q_{\mathrm{NL}} = \begin{pmatrix}
A_1^{(0)} & A_0 & & & \\
A_2 & A_1 & A_0 & & \\
 & A_2 & A_1 & A_0 & \\
 & & \ddots & \ddots & \ddots \\
 & & & &
\end{pmatrix}
$$

with

$$
A_1^{(0)} = \left(\begin{array}{c|cccc}
 & 0 & 1 & 2 & 3 \\
\hline
0 & -\lambda_{\mathrm{NL}} & & & \\
1 & \nu & -(\nu + \lambda_{\mathrm{NL}}) & & \\
2 & & \nu & -(\nu + \lambda_{\mathrm{NL}}) & \\
3 & & & \nu & -(\nu + \lambda_{\mathrm{NL}})
\end{array}\right),
$$

$$
A_0 = \left(\begin{array}{c|cccc}
 & 0 & 1 & 2 & 3 \\
\hline
0 & \lambda_{\mathrm{NL}} & & & \\
1 & & \lambda_{\mathrm{NL}} & & \\
2 & & & \lambda_{\mathrm{NL}} & \\
3 & & & & \lambda_{\mathrm{NL}}
\end{array}\right), \qquad
A_2 = \left(\begin{array}{c|cccc}
 & 0 & 1 & 2 & 3 \\
\hline
0 & & \mu & & \\
1 & & & \mu & \\
2 & & & & \mu \\
3 & & & &
\end{array}\right),
$$

$$A_1 = \begin{pmatrix} & 0 & 1 & 2 & 3 \\ \hline 0 & -(\mu+\lambda_{\text{NL}}) & & & \\ 1 & \nu & -(\mu+\nu+\lambda_{\text{NL}}) & & \\ 2 & & \nu & -(\mu+\nu+\lambda_{\text{NL}}) & \\ 3 & & & \nu & -(\nu+\lambda_{\text{NL}}) \end{pmatrix}.$$

Using parameters K, K_W, R and V from Definition 2.1.40 we have $K := \{0,1,2,3\}$,

$$K_W := \{0,1,2\},\ R := \begin{pmatrix} & 0 & 1 & 2 & 3 \\ \hline 0 & & 1 & & \\ 1 & & & 1 & \\ 2 & & & & 1 \\ 3 & & & & \end{pmatrix},\ \text{and } V := \begin{pmatrix} & 0 & 1 & 2 & 3 \\ \hline 0 & & & & \\ 1 & \nu & -\nu & & \\ 2 & & \nu & -\nu & \\ 3 & & & \nu & -\nu \end{pmatrix}.$$

Following Section 2.1.4.2, we construct a loss system with the same parameters K, K_W, μ, R and V as in Definition 2.2.34, and with λ adjusted in such a way, that the loss system and the system without loss have the same throughput λ_{NL}.

Definition 2.2.35 (Crusher station with loss)**.** Given a tandem system with two nodes. The service times at both nodes are exponential. The stochastic process $(X(t) : t \geq 0)$ with $X(t) \in \mathbb{N}_0$ describes the number of truck loads at the first node. The stochastic process $(Y(t) : t \geq 0)$ with $Y(t) \in K := \{0,1,2,3\}$ describes number of truck loads in the second node. The first node models a discharging platform with service rate μ and infinite number of waiting places. The second node models a crusher with service rate ν with two additional waiting places in a bin. It can contain maximal three truck loads. The truck loads arrive in a Poisson stream with rate λ at the first node such that the throughput in steady state is λ_{NL} from Definition 2.2.34:

$$\lambda := \lambda \sum_{n\in\mathbb{N}_0} \sum_{k\in\{0,1,2\}} \pi(n,k) = \lambda_{\text{NL}}$$

where $\pi(n,k)$ is the steady-state distribution of the loss system. When the second node is full, the first node interrupts service and newly arriving truck loads to the first node are *lost*.

The generator of the system $Q = (q((n,k),(n',k')) : n,n' \in \mathbb{N}, k,k' \in K)$ is

$$\begin{aligned} q((n,k),(n+1,k)) &= \lambda, && k \in \{0,1,2\}, n \in \mathbb{N}_0, \\ q((n,k),(n-1,k+1)) &= \mu, && k \in \{0,1,2\}, n > 0, \\ q((n,k),(n,k-1)) &= v \in \mathbb{R}_0^+, && k \in \{1,2,3\}, n \in \mathbb{N}_0, \\ q((n,k),(i,m)) &= 0, && \text{otherwise for } (n,k) \neq (i,m). \end{aligned}$$

In order to test, weather the additional computation effort for adjusting of λ is worthwhile, we also construct a very simple model where we completely ignore the blocking problem and that both queues may influence each other. The resulting simple model consists of independent $M/M/1/\infty$ and $M/M/1/2$ queues.

Definition 2.2.36 (Crusher station's subsystems as independent queues)**.** Given a system with two dependent nodes. The service times at both nodes are exponential. The stochastic process $(X_{\text{simple}}(t) : t \geq 0)$ with $X_{\text{simple}}(t) \in \mathbb{N}_0$ describes the number of truck loads at the first node. The stochastic process $(Y_{\text{simple}}(t) : t \geq 0)$ with $Y_{\text{simple}}(t) \in K := \{0, 1, 2, 3\}$ describes number of truck loads in the second node. The first node models a discharging platform with service rate μ and infinite number of waiting places. The second node models a crusher with service rate ν with two additional waiting places in a bin. It can contain maximal three truck loads. If the second node is full the new arriving truck loads *to the second node* are lost. The arrival rate at both nodes is λ_{NL}. See Figure 2.2.10 on page 83.

The generator of the system $Q_{\text{simple}} = (q_{\text{simple}}((n, k), (n', k')) : n, n' \in \mathbb{N}, k, k' \in K)$ is

$$\begin{aligned}
q_{\text{simple}}((n, k), (n + 1, k)) &= \lambda_{NL}, && k \in \{0, 1, 2, 3\}, n \in \mathbb{N}_0, \\
q_{\text{simple}}((n, k), (n - 1, k)) &= \mu, && k \in \{0, 1, 2, 3\}, n > 0, \\
q_{\text{simple}}((n, k), (n, k + 1)) &= \lambda_{NL}, && k \in \{0, 1, 2\}, n \in \mathbb{N}_0, \\
q_{\text{simple}}((n, k), (n, k - 1)) &= v \in \mathbb{R}_0^+, && k \in \{1, 2, 3\}, n \in \mathbb{N}_0, \\
q_{\text{simple}}((n, k), (i, m)) &= 0, && \text{otherwise for } (n, k) \neq (i, m).
\end{aligned}$$

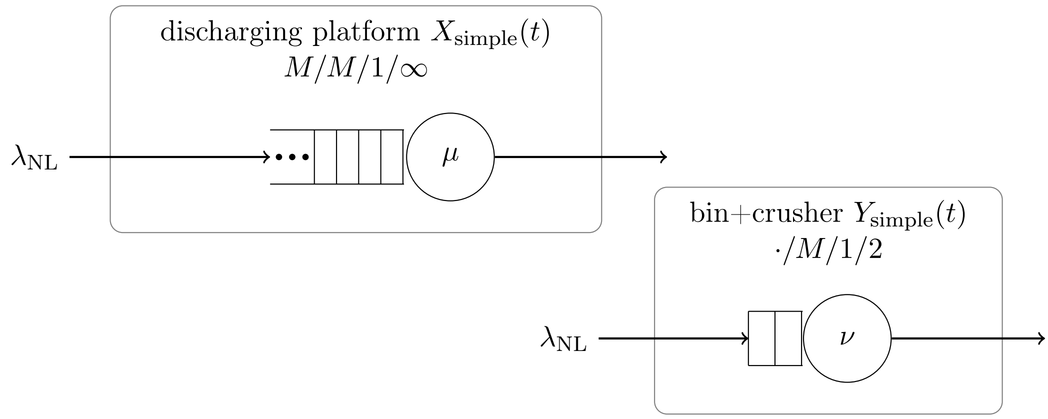


Figure 2.2.10.: Discharging point and crusher as a simple model from Definition 2.2.36.

Corollary 2.2.37. *The system without loss defined in Definition 2.2.34 is ergodic if and only if*

$$\mu \frac{\nu^3 + \nu^2\mu + \nu\mu^2}{\nu^3 + \nu^2\mu + \nu\mu^2 + \mu^3} > \lambda_{NL}. \tag{2.2.50}$$

An appropriate ergodic loss system with the same effective arrival rate λ_{NL} as defined in Definition 2.2.35 exists if and only if the system without loss is ergodic. The loss system approximation is unique.

For the ergodic loss system holds:

- *Stationary distribution π is*

$$\pi(n, k) = \left(1 - \frac{\lambda}{\mu}\right) \frac{\lambda^n}{\mu^n} \frac{\nu^3}{\nu^3 + \nu^2\lambda + \nu\lambda^2 + \lambda^3} \left(\frac{\lambda}{\nu}\right)^k. \tag{2.2.51}$$

- *The mean number of truck loads at discharging station in steady state is*

$$E(X) = \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}} = \frac{\lambda}{\mu - \lambda}.$$

- *The mean number of truck loads at crusher in steady state is*

$$E(Y) = \frac{\nu^2\lambda + 2\nu\lambda^2 + 3\lambda^3}{\nu^3 + \nu^2\lambda + \nu\lambda^2 + \lambda^3}.$$

Proof. The loss system approximation of the crusher station model is a special case of a tandem network defined in Section 2.2.3. So we can obtain $f(x)$ just by replacing λ through x in θ in (2.2.37)

$$G = 1 + \frac{x}{\nu} + \frac{x^2}{\nu^2} + \frac{x^3}{\nu^3} = \frac{\nu^3 + \nu^2 x + \nu x^2 + x^3}{\nu^3},$$

$$f(x)(k) = G^{-1}\left(\frac{x}{\nu}\right)^k = \frac{\nu^3}{\nu^3 + \nu^2 x + \nu x^2 + x^3}\left(\frac{x}{\nu}\right)^k.$$

It also holds

$$x \sum_{k \in K_W} f(x)(k) = x\big(1 - f(x)(3)\big) = x\frac{\nu^3 + \nu^2 x + \nu x^2}{\nu^3 + \nu^2 x + \nu x^2 + x^3}$$

and

$$\frac{\partial}{\partial x}\left(x \sum_{k \in K_W} f(x)(k)\right) = \frac{v^4\left(\nu^2 + 2\nu x + 3x^2\right)}{\left(\nu^3 + \nu^2 x + \nu x^2 + x^3\right)} > 0.$$

According to Proposition 2.1.45 the tandem system without loss, as defined in Definition 2.2.34, is positive recurrent if and only if

$$\mu \cdot \big(1 - f(\mu)(3)\big) \quad > \quad \lambda_{\text{NL}}.$$

This is equivalent to (2.2.50).

Because $\frac{\partial}{\partial x}\big(x \sum_{k \in K_W} f(x)(k)\big)$ is positive for all x, the function $x \sum_{k \in K_W} f(x)(k)$ is monotone increasing in x. We can apply Proposition 2.1.47 which states that the system without loss is ergodic if and only if the corresponding loss system with the same effective arrival rate is ergodic.

Substituting $\xi(n) = \left(1 - \frac{\lambda}{\mu}\right)\frac{\lambda^n}{\mu^n}$ and $\theta = f(\lambda)$ in $\pi(n,k) = \xi(n)\theta(k)$ we obtain (2.2.51). The mean number of truck loads in the loss system corresponds to the mean number of customers an $M/M/1/\infty$ queue with arrival rate λ and service rate μ. It is $\frac{\lambda}{\mu - \lambda}$. The mean number of truck loads at the crusher is $\theta(1) + 2\theta(2) + 3\theta(3)$. □

	capacity of the bin N						
	0	1	2	3	4	5	∞
$\lambda_{\mathrm{NL,critical}}$	24.877	31.314	33.561	34.424	34.767	34.906	$\nu = 35$

Table 2.1.: $\lambda_{\mathrm{NL,critical}}$

2.2.6.2. Evaluation of loss-system approximation

In the following example we will compare a particular system without loss with its loss-system approximation. The parameters are derived from parameters of a real world system analyzed in [ZW09]. The steady-state distribution of the system without loss is calculated numerically with matrix-geometric methods.

Example 2.2.38. Consider a system without loss as defined in Definition 2.2.34, with $\mu = 2 \cdot \mu_{\mathrm{orgn}} = 2 \cdot 43$ and $\nu = 35$. The service rate $\mu = 2 \cdot 43$ approximates service rates of the original $M/M/2/\infty$ system in [ZW09] where each server has service rate $\mu_{\mathrm{orgn}} = 43$. The service rate ν corresponds to the value which the authors in [ZW09] call "throughput of the crusher" or "crusher's crushing capacity". For us the most plausible interpretation of this value is the service rate.

We calculate critical arrival rate $\lambda_{\mathrm{NL,critical}}$. That means, a system without loss with arrival rate λ_{NL} is positive recurrent if and only if $\lambda_{\mathrm{NL}} < \lambda_{\mathrm{NL,critical}}$. Using the formula (2.2.50) with ">" replaced by "=" and λ_{NL} replaced by $\lambda_{\mathrm{NL,critical}}$ we calculate

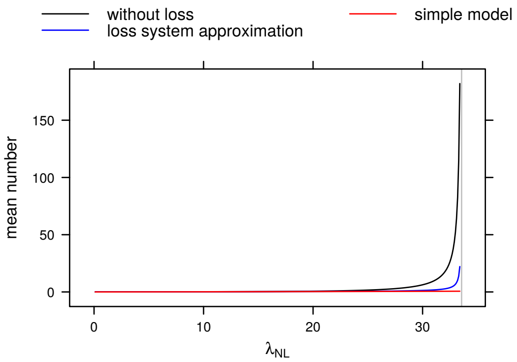
$$\lambda_{\mathrm{NL,critical}} \approx 33.56144$$

We also compute input rates $\lambda_{\mathrm{NL,critical}}$ for different capacity of the bin at the crusher. See Table 2.1 on page 85.

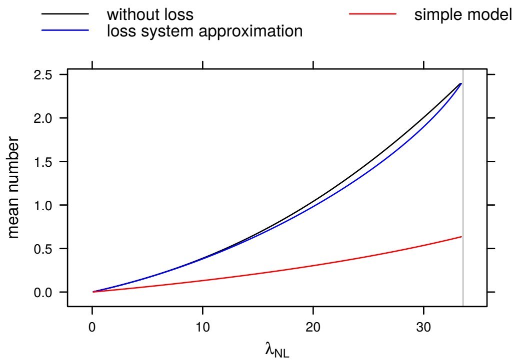
We see that $\lambda_{\mathrm{NL,critical}} \approx 33.56144$ with bin capacity of two loads is pretty close to the maximal throughput ν. Under assumption that our simplified tandem with the nodes of type $M/M/1/\infty$ and $\cdot/M/1/2$ behaves similarly to the original system from [ZW09] with $M/M/2/\infty$ and $\cdot/D/1/2$ nodes, the choice of the bin capacity of only two loads for the original system is optimal.

The plots in Figure 2.2.11 on page 86 show a steep increase of the mean truck loads number of both systems when they approach the critical arrival rate $\lambda_{\mathrm{NL,critical}}$. The plots in Figure 2.2.12 on page 87 show zoomed in section of data of the same systems when the average number of customers is not large.

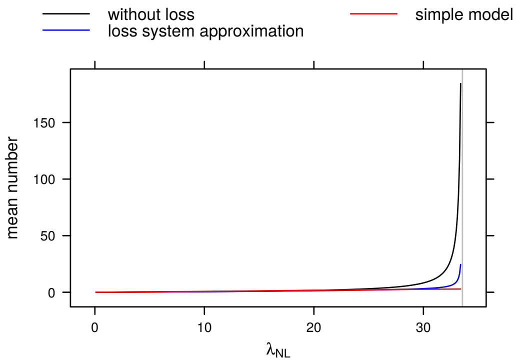
We see that the loss-system approximation from Definition 2.2.35, constructed according to Section 2.1.4, in general does not approximate well the average number of loads at the discharging platform and total number of loads of the system without loss defined in Definition 2.2.34, despite same service rates and effective arrival rates in both systems. There is a small interval – approximately $[0, 10]$ – where the loss system approximation is quite good. This good approximation results were expected according to Section 2.1.4.3 due to small blocking probability, see Figure 2.2.13a on page 88. But, we also need to take into account that on interval $[0, 10]$ this approximation is nearly as good as the simple model with independent queues from Definition 2.2.36. So we can conclude that in this particular example the loss system approximation is not worth for estimation of mean number of truck loads at the discharging point and in the whole system.



(a) Mean value of truck loads at discharging station.



(b) Mean value of truck loads at crusher.



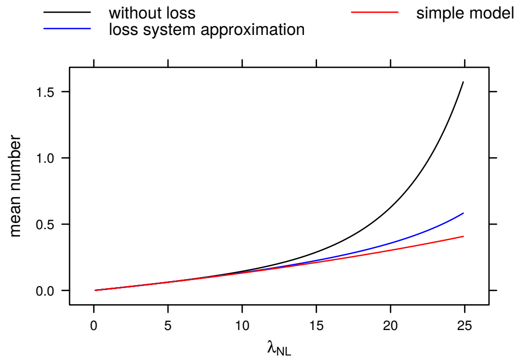
(c) Mean number of truck loads in the system.

Figure 2.2.11.: Mean number of truck loads in crusher station with $\mu = 2 \cdot 43$, $\nu = 35$. Gray line indicates the critical value of λ_{NL}. At and on the right of this gray line the system without loss and the approximating loss system are not more stable.

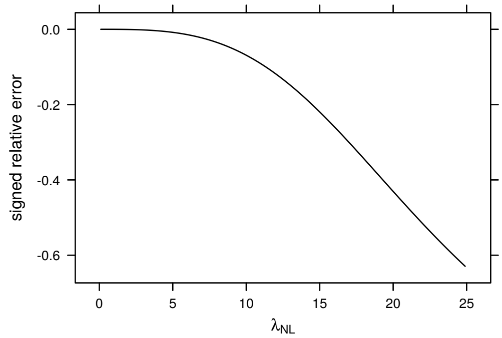
We also cannot see sharp increasing of the mean number of trucks in the queue after the average truck arrival rate reaches 28 as it is stated in [ZW09, p. 3]. Nevertheless both systems become instable with the same parameter λ_{NL} as it was proved in Corollary 2.2.37.

We see that the average number of loads at the crasher is well approximated, but their contribution to the total number of loads is negligible.

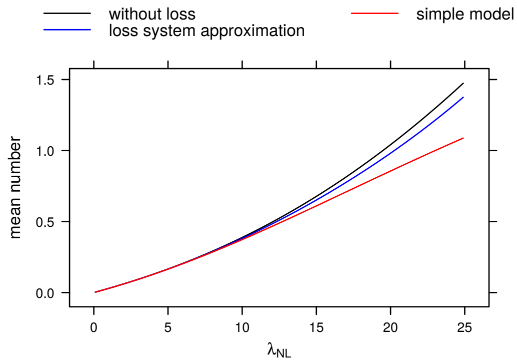
An interesting property is that the mean number of customers in Figure 2.2.11 on page 86 and in Figure 2.2.12 on page 87 of the loss-system approximation is always lower than the mean number of customers in the system without loss. This behavior can be a topic for future research, because this may indicate that the loss-system approximation generates strict lower bounds for systems without loss.



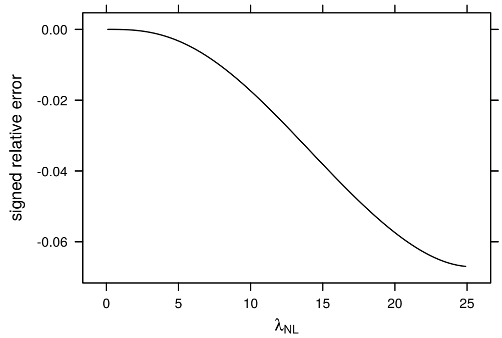
(a) Mean value of truck loads at discharging station.



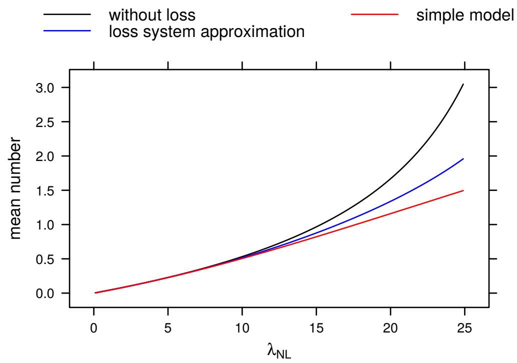
(b) Signed relative error of mean value of truck loads at discharging station $\frac{E(X)-E(X_{\mathrm{NL}})}{E(X_{\mathrm{NL}})}$.



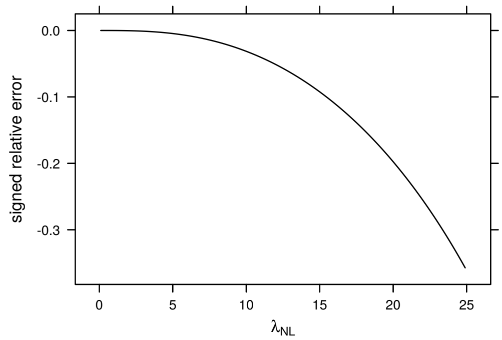
(c) Mean value of truck loads at crusher.



(d) Signed relative error of mean value of truck loads at crusher $\frac{E(Y)-E(Y_{\mathrm{NL}})}{E(Y_{\mathrm{NL}})}$.

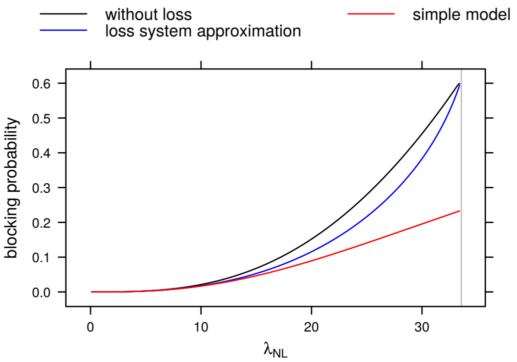


(e) Mean number of truck loads in the system.

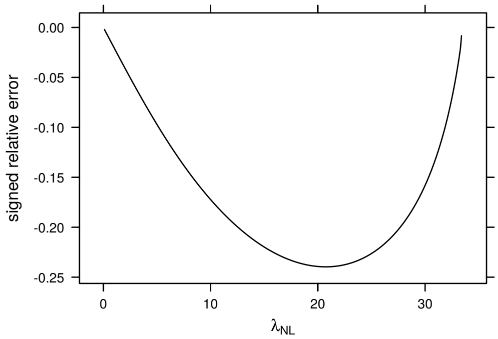


(f) Signed relative error of mean number of truck loads in the system $\frac{E(X+Y)-E(X_{\mathrm{NL}}+Y_{\mathrm{NL}})}{E(X_{\mathrm{NL}}+Y_{\mathrm{NL}})}$.

Figure 2.2.12.: Mean number of truck loads in crusher station with $\mu = 2 \cdot 43$, $\nu = 35$ only region for $\lambda_{\mathrm{NL}} \leq 25$ and the corresponding signed relative errors.



(a) Blocking probability for system, i.e. $P(Y_{\mathrm{NL}} = 3)$.



(b) Signed relative error of blocking probability $\frac{P(Y=3)-P(Y_{\mathrm{NL}}=3)}{P(Y_{\mathrm{NL}}=3)}$.

Figure 2.2.13.: Blocking probability for crusher station with $\mu = 2 \cdot 43$, $\nu = 35$ and the corresponding signed relative errors.

3. Embedded Markov chains analysis

We continue our investigations of single-queue systems in a random environment from Section 2. The systems live in continuous time and we now observe them at departure instants only, which results in considering an embedded Markov chain. We find that the behaviour of the embedded Markov chain is often considerably different from that of the original continuous time Markov process investigated in Section 2.

Our aim is to identify conditions which guarantee that even for the embedded Markov chain a product form equilibrium exists at the discrete observation time points as well.

For exponential queueing systems we show that there is a product form equilibrium under rather general conditions. For systems with non-exponential service times more restrictive constraints are needed, which we prove by a counter example where the environment represents an inventory attached to an $M/D/1$ queue. Such integrated queueing-inventory systems are dealt with in the literature previously, see [KLM11]. Further applications are, e.g., in modeling unreliable queues.

For investigating $M/G/1/\infty$ queues embedded Markov chains provide a standard procedure to avoid using supplementary variable technique. Embedded chain analysis was applied by Vineetha [Vin08] who extended the theory of integrated queueing-inventory models with exponential service times to systems with service times which are i.i.d. and follow a general distribution. Our investigations which are reported in this section were in part motivated by hers.

In Section 3.2 we revisit some of Vineetha's [Vin08] queueing-inventory systems, using similarly embedded Markov chain techniques. In the course of these investigations we found that there arise problems even for purely exponential systems, which we describe in Section 3.1.1 and Section 3.1.2 first, before describing the $M/G/1/\infty$ queue in a random environment and its structural properties.

To emphasize the problems arising from the interaction of the two components of integrated systems, we remind the reader, that for ergodic $M/G/1/\infty$ queues the limiting and stationary distribution of the continuous time queue length process and the Markov chains embedded at departure instants are the same.

Our first finding is, that even in the case of $M/M/1/\infty$ queues with attached inventory this in general does not hold. This especially implies, that the product form results obtained in Section 2 do not carry over immediately to the case of loss systems in a random environment observed at departure times from the queue (downward jumps of the generalized birth-death process).

A striking observation is that for a system which is ergodic in the continuous time Markovian description the Markov chain embedded at departure instants may be not ergodic. The reason for this is two-fold. Firstly, the embedded Markov chain may have inessential states due to the specified interaction rules. Secondly, even when we delete all inessential states, the resulting single positive recurrent class may be periodic.

We study this problem in depth in Section 3.1.2 for purely exponential systems, and provide a set of examples which elucidate the problems which one is faced with. Our main result in this section proves the existence of a product form steady-state distribution (which is not necessary a limiting distribution) for the Markov chain embedded at departure instants and provides a precise connection between the steady states of the continuous time process and the embedded chain (Theorem 3.1.16).

It turns out, that a similar result in the setting with $M/G/1/\infty$ queues is not valid. We are able to give sufficient conditions for the structure of the environment, which guarantee the existence of product form equilibria (Theorem 3.2.25).

Unfortunately enough, an analogue to Theorem 3.1.16 is not valid for systems with non exponential service times. We prove this by constructing a counterexample which is an $M/D/1/\infty$ queue with an attached environment in Proposition 3.2.8 in Section 3.2.

3.1. M/M/1/∞ queueing system in a random environment

Recall, that the paths of Z are cadlag. With $\tau_0 = \sigma_0 = \zeta_0 = 0$ and

$$\tau_{n+1} := \inf(t > \tau_n : X(t) < X(t^-)), \quad n \in \mathbb{N},$$

denote the sequence of departure times of customers by $\tau = (\tau_0, \tau_1, \tau_2, \dots)$, and with

$$\sigma_{n+1} := \inf(t > \sigma_n : X(t) > X(t^-)), \quad n \in \mathbb{N},$$

denote by $\sigma = (\sigma_0, \sigma_1, \sigma_2, \dots)$ the sequence of instants when arrivals are admitted to the system (because the environment is in states of K_W, i.e., not blocking)
and with

$$\zeta_{n+1} := \inf(t > \zeta_n : Z(t) \neq Z(\zeta_n)), \quad n \in \mathbb{N},$$

denote by $\zeta = (\zeta_0, \zeta_1, \zeta_2, \dots)$ the sequence of jump times of Z.

The following lemma will be used in the sequel. It refers to the structure of the continuous time process. We emphasize that the generator V is not necessarily irreducible.

Lemma 3.1.1. *For any strictly positive $\eta \in \mathbb{R}^+$ the matrix $(-\operatorname{diag}(V) + \eta I_W)$ is invertible.*

Proof. For any $k \in K_W$ the corresponding diagonal element of the matrix $(-\operatorname{diag}(V) + \eta I_W)$ is greater than η because $-v(k,k) \geq 0$.

If $k \in K_B$, we utilize the ergodicity of Z in continuous time and apply Lemma A.2.1 with $\widetilde{K}_B := \{k\}$. The lemma implies that there is some $m \neq k$ with $v(k,m) > 0$. It follows $-v(k,k) > 0$.

We conclude that the diagonal matrix $(-\operatorname{diag}(V) + \eta I_W)$ has only strictly positive values on its diagonal and therefore it is invertible. □

3.1.1. Observing the system at departure instants

Recall that the paths of Z are cadlag and that $\tau = (\tau_0, \tau_1, \tau_2, \dots)$ with $\tau_0 = 0$ denotes the sequence of departure times of customers. Then with $\hat{X}(n) := X(\tau_n)$ and $\hat{Y}(n) := Y(\tau_n)$ for $n \in \mathbb{N}_0$ it is easy to see that

$$\hat{Z} = ((\hat{X}(n), \hat{Y}(n)) : n \in \mathbb{N}_0) \tag{3.1.1}$$

is a homogeneous Markov chain on state space $E = \mathbb{N}_0 \times K$. If $\hat{Z}$ has a unique stationary distribution, this will be denoted by $\hat{\pi}$.

Definition 3.1.2. If the embedded Markov chain $\left((\hat{X}(n), \hat{Y}(n)) : n \in \mathbb{N}_0\right)$ of the loss system in a random environment with state space $E := \mathbb{N}_0 \times K$ has a unique steady-state distribution, we denote this distribution

$$\hat{\pi} \quad := \quad (\hat{\pi}(n,k) : (n,k) \in E := \mathbb{N}_0 \times K).$$

and the marginal steady-state distributions

$$\hat{\xi} \quad := \quad (\hat{\xi}(n) : n \in \mathbb{N}_0), \qquad \hat{\xi}(n) := \sum_{k \in K} \hat{\pi}(n,k),$$

$$\hat{\theta} \quad := \quad (\hat{\theta}(k) : k \in K), \qquad \hat{\theta}(k) := \sum_{n \in \mathbb{N}_0} \hat{\pi}(n,k).$$

It will turn out that this Markov chain exhibits interesting structural properties of the loss systems in random environments. E.g., with ξ from (2.1.8) we will prove that

$$\hat{\pi}(n,k) = \hat{\xi}(n) \cdot \hat{\theta}(k) \text{ with } \hat{\xi}(n) = \xi(n)$$

holds, but in general we do not have $\hat{\pi}(n,k) > 0$ on the global state space E, because $\hat{\theta}(k) = 0$ may occur. Especially, in general it holds $\theta \neq \hat{\theta}$.

The reason for this seems to be the rather general vice-versa interaction of the queueing system and the environment. Of special importance is the fact that we consider the continuous time systems at departure instants where we have the additional information that right now the influence of the queueing systems on the change of the environment is in force (described by the stochastic matrix R).

The dynamics of $\hat{Z}$ will be described in a way that resembles the $M/G/1$ type matrix analytical models. Recall that the state space E carries an order structure, see Remark 2.1.4, which will govern the description of the transition matrix and, later on, of the steady-state vector.

Definition 3.1.3. We define the one-step transition matrix $\mathbf{P}$ of $\hat{Z}$ by

$$\begin{aligned}&\left(\mathbf{P}_{(i,k),(j,m)} : (i,k),(j,m) \in E\right)\\ &:= \left(P(Z(\tau_1) = (j,m) | Z(0) = (i,k)) : (i,k),(j,m) \in E\right),\end{aligned}$$

and introducing matrices $A^{(i,n)} \in \mathbb{R}^{K \times K}$ and $B^{(n)} \in \mathbb{R}^{K \times K}$ by

$$B_{km}^{(n)} := P(Z(\tau_1) = (n,m) | Z(0) = (0,k)) \tag{3.1.2}$$

$$A_{km}^{(i,n)} := P(Z(\tau_1) = (i+n-1,m) | Z(0) = (i,k)), \qquad 1 \leq i, \tag{3.1.3}$$

for $k, m \in K$, the matrix $\mathbf{P}$ has the form

$$\mathbf{P} = \begin{pmatrix} B^{(0)} & B^{(1)} & B^{(2)} & B^{(3)} & \dots \\ A^{(1,0)} & A^{(1,1)} & A^{(1,2)} & A^{(1,3)} & \dots \\ 0 & A^{(2,0)} & A^{(2,1)} & A^{(2,2)} & \dots \\ 0 & 0 & A^{(3,0)} & A^{(3,1)} & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}, \tag{3.1.4}$$

which exploits the structure of the state space as a product of level variables in $\mathbb{N}_0$ and phase variables in K.

For the loss system in a random environment we will solve the equation

$$\hat{\pi}\mathbf{P} = \hat{\pi} \tag{3.1.5}$$

for a stochastic solution $\hat{\pi}$ which is a steady-state distribution of the embedded Markov chain $\hat{Z}$. Because $\hat{Z}$ is in general not irreducible on E there are some subtleties with respect to the uniqueness of a stochastic solution of this equation.

For further calculations it will be convenient to group $\hat{\pi}$ according to the queue length:

Definition 3.1.4. We write $\hat{\pi}$ as

$$\hat{\pi} = (\hat{\pi}^{(0)}, \hat{\pi}^{(1)}, \hat{\pi}^{(2)}, \dots) \tag{3.1.6}$$

with

$$\hat{\pi}^{(n)} = (\hat{\pi}(n,k) : k \in K), \quad n \in \mathbb{N}_0. \tag{3.1.7}$$

Especially we write for $(n,k) \in E$

$$\hat{\pi}^{(n)}(k) := \hat{\pi}(n,k).$$

An immediate consequence of this definition is that the steady-state equation (3.1.5) can be written as

$$\hat{\pi}^{(0)} B^{(n)} + \sum_{i=1}^{n+1} \hat{\pi}^{(i)} A^{(i,n-i+1)} = \hat{\pi}^{(n)}, \qquad n \in \mathbb{N}_0. \tag{3.1.8}$$

3.1.2. Steady state for the system observed at departure instants

The main idea in this section is to show that $\hat{\pi}(n,k) = \hat{\xi}(n)\hat{\theta}(k)$ with $\hat{\xi} = \xi$, $\hat{\theta} = c^{-1}\theta I_W R$ and normalization constant c solves the steady-state equation (3.1.5). The expression $\hat{\theta} = c^{-1}\theta I_W R$ originally comes from our previous investigations, where in [KD13b] we directly computed $\hat{\theta}$ under the assumption that $(\lambda I_W - V)^{-1}$ exists. Later on we proved that $\hat{\theta} = c^{-1}\theta I_W R$.

An explanation for the special form of $\hat{\theta}$ was given to us by an anonymous reviewer in a report on a previous version of this section submitted to the journal *Queuing Systems.* The text within angle $\langle\rangle$ braces was adjusted to present references and bibliography keys:

> "The process $(X(t), Y(t))$ is a (level-dependent) QBD process, as is also mentioned by the authors, with a particular structure in which transitions from level n to level $n+1$, n and $n-1$, are governed by matrices $A_0 = \lambda I_W$, $A_1(n) = V - A_0 - \mathrm{diag}(A_2(n)\mathbf{e})$, and $A_2(n) = \mu(n) I_W R$ respectively.

So the stationary distribution embedded at service completions is just the distribution of the QBD at hand upon moving down one level, which is easily derived from π. In particular we have

$$\hat{\pi}(n,k) \equiv P(\hat{X}=n,\hat{Y}=k) = c^{-1}\sum_{\ell\in K} P(X=n+1, Y=\ell)A_2(n+1)_{\ell,k} \tag{3.1.9}$$

here c follows from normalization. Using the product-form for π as found in ⟨[KD12]⟩, namely $\pi(n,k)=\xi(n)\theta(k)$, see ⟨(2.1.4)⟩, this leads to

$$\hat{\pi}(n,k) = c^{-1}\xi(n+1)\mu(n+1)\sum_{\ell\in K}\theta(\ell)[I_W R]_{\ell,k}$$

which reduces to $c^{-1}\lambda\xi(n)\sum_{\ell\in K}\theta(\ell)[I_W R]_{\ell,k}$. This proves the main statements in ⟨Theorem 3.1.16 (a)⟩, namely the product form in ⟨(3.1.39)⟩ together with the fact ⟨(3.1.40)⟩ that $\hat{\xi}(n)=\xi(n)$, and the particular form of $\theta(k)$ in ⟨(3.1.41)⟩ . The result ⟨(3.1.38) in Proposition 3.1.14⟩ then follows immediately from ⟨(3.1.41)⟩ ($\theta_k = 0$ when $k\notin L$)."

The formula (3.1.9) provided by the reviewer points to an other intuitive explanation for the special form of the $\hat{\pi}$. Let $Z(t)=(X(t),Y(t))$ be an ergodic stationary process with generator Q and $\hat{Z}=(\hat{X},\hat{Y})$ be the corresponding embedded chain at departure times. The average number of jumps to state (n,k) directly from level $n+1$ is

$$\begin{aligned}
&E[\text{Number of jumps into } (n,k) \text{ directly from level } n+1]\\
&= \mu(n+1)\sum_{\ell\in K_W}\pi(n+1,\ell)R_{\ell,k}\\
&= \underbrace{\xi(n+1)}_{=\xi(n)\lambda/\mu(n+1)}\ \mu(n+1)\sum_{\ell\in K_W}\theta(\ell)[I_W R]_{\ell,k}\\
&= \lambda\xi(n)\sum_{\ell\in K_W}\theta(\ell)[I_W R]_{\ell,k} + \lambda\xi(n)\sum_{\ell\in K_B}\underbrace{\theta(\ell)[I_W R]_{\ell,k}}_{=0 \text{ for } \ell\in K_B}\\
&= \lambda\xi(n)\sum_{\ell\in K}\theta(\ell)[I_W R]_{\ell,k}.
\end{aligned}$$

The probability $\hat{\pi}(n,k)$ is the proportion of the jumps into (n,k) directly from level $n+1$ in unit of time

$$\begin{aligned}
\hat{\pi}(n,k) &= \frac{E[\text{Number of jumps into } (n,k) \text{ directly from level } n+1]}{\sum_{(i,m)\in\mathbb{N}_0\times K} E[\text{Number of jumps into } (i,m) \text{ directly from level } i+1]}\\
&= \frac{\lambda\xi(n)\sum_{\ell\in K}\theta(\ell)[I_W R]_{\ell,k}}{\lambda\underbrace{\sum_{i\in\mathbb{N}}\xi(i)}_{=1}\underbrace{\sum_{m\in K}\sum_{\ell\in K}\theta(\ell)[I_W R]_{\ell,m}}_{=:c}} = \xi(n)\,\underbrace{c^{-1}\theta(k)\sum_{\ell\in K_W}[I_W R]_{\ell,k}}_{=:\hat{\theta}(k)}\,.
\end{aligned}$$

Both ideas for the special form of $\hat{\theta}$ are not rigorous proofs and still require a careful construction of all involved stochastic processes. Nevertheless they deliver interesting insights to the embedded chains. The main advantages of both approaches are the simplicity of the solution as soon as their formula is mathematically justified, they do not require the matrix $\mathbf{P}$ from (3.1.4), and they do not require the special assumption about invertibility of the matrix $(\lambda I_W - V)$. The disadvantage of both approaches is that it requires the knowledge of the steady-state solution $\pi(n,k)$ in continuous time, which is not alway available. In our version of the proof we will use a hybrid approach: we will analyze the matrix $\mathbf{P}$ and show that the proposed $\hat{\pi}(n,k) = \xi(n)\hat{\theta}(k)$ with $\hat{\theta} = c^{-1}\theta I_W R$ solves the steady-state equation $\hat{\pi}\mathbf{P} = \hat{\pi}$. The reasons for us to favor the hybrid approach are

1. The analysis of the matrix $\mathbf{P}$ provides insight into the embedded processes and we will use it to calculate steady-state distribution for some $M/G/1/\infty$ systems where the continuous time probability $\pi(n,k)$ is not available.
2. We do not need to assume about invertibility of $(\lambda I_W - V)$ in contrast to the direct approach in [KD13b].

Assumption 3.1.5 (Overall assumption for Section 3). *In this section we will frequently use matrix operations for matrices from $\mathbb{R}^{K\times K}$, where K is the environment state space. We assume that matrix multiplication is associative. When K is finite, the associativity of matrix multiplication is granted, but when K is infinite, problems may occur. The reason for these problems is the fact, that in the case of infinite K, each matrix multiplication creates a new matrix whose elements are infinite series. The value of these series can depend on the summation order.*

In a first step we analyze the dynamics incorporated in the matrix $A^{(i,n)}$ and $B^{(n)}$.

We start our investigation with a detailed analysis of the one-step transition matrix (3.1.4) and will express the matrices $B^{(n)}$ and $A^{(i,n)}$ from Definition 3.1.3 by means of auxiliary matrices W and $U^{(i,n)}$, which reflect the dynamics of the system.

Lemma 3.1.6. *Recall that τ_1 denotes the first departure instant, that σ_1 denotes the first arrival instant of an admitted customer, and that $Y(\sigma_1) \in K_W$ holds.*

For $k \in K, m \in K_W$, we define the matrix $U^{(i,n)}$, whose entries $U^{(i,n)}_{km}$ are probabilities for the system starting with i customers and environment state k to admit n new customers and being in environment state m right before the first service is finished.

$$U^{(i,n)}_{km} := P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, m) | Z(0) = (i,k)\right), \qquad 1 \le i, n \in \mathbb{N}_0, \tag{3.1.10}$$

and for $k \in K$ and $m \in K_B$ we prescribe by definition $U^{(i,n)}_{km} = 0$.

Similarly, for $k \in K, m \in K_W$, we define the matrix W, whose entries W_{km} are probabilities for the empty system starting in environment state k to be in environment state m when the first customer is admitted to the system.

$$W_{km} := P(Z(\sigma_1) = (1,m) | Z(0) = (0,k)), \tag{3.1.11}$$

and for $k \in K$ and $m \in K_B$ prescribe by definition $W_{km} = 0$.

Then it holds for $A^{(i,n)}$ and $B^{(n)}$ from Definition 3.1.3

$$A^{(i,n)} = U^{(i,n)} R, \tag{3.1.12}$$

$$B^{(n)} = W A^{(1,n)} = W U^{(1,n)} R. \tag{3.1.13}$$

Proof. Using the fact, that the paths of the system in continuous time almost sure have left limits, we get for $i > 0,\ n \geq 0$ and $k, m \in K$

$$\begin{aligned}
A_{km}^{(i,n)} &= P\left((X(\tau_1), Y(\tau_1)) = (i+n-1, m) | Z(0) = (i,k)\right) \\
&= \sum_{h\in K} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, h) | Z(0) = (i,k)\right) R(h,m) \\
&= \sum_{h\in K} U_{kh} R(h,m),
\end{aligned}$$

which in matrix form is (3.1.12).

For the property (3.1.13) we will use the fact, that if the system starts with an empty queue, then the first arrival occurs always before the first departure, $P(\sigma_1 < \tau_1) = 1$. We will also use the strong Markov (SM) property of Z. We obtain for $n \geq 0$ and $k, m \in K$

$$\begin{aligned}
B_{km}^{(n)} :=& P((X(\tau_1), Y(\tau_1)) = (n, m) | Z(0) = (0,k)) \\
=& \sum_{h\in K} P\left((X(\tau_1), Y(\tau_1)) = (n, m) \cap Z(\sigma_1) = (1,h) | Z(0) = (0,k)\right) \\
=& \sum_{h\in K} P\left((X(\tau_1), Y(\tau_1)) = (1+n-1, m) | Z(\sigma_1) = (1,h) \cap Z(0) = (0,k)\right) \\
& \cdot P\left(Z(\sigma_1) = (1,h) | Z(0) = (0,k)\right) \\
\overset{SM}{=}& \sum_{h\in K} \underbrace{P\left((X(\tau_1), Y(\tau_1)) = (1+n-1, m) | Z(0) = (1,h)\right)}_{=A_{hm}^{(1,n)}} \\
& \cdot \underbrace{P\left(Z(\sigma_1) = (1,h) | Z(0) = (0,k)\right)}_{=W_{kh}} \\
=& \sum_{h\in K} W_{kh} A_{hm}^{(1,n)},
\end{aligned}$$

which proves (3.1.13). □

The proof of Lemma 3.1.6 reveals that the stochastic matrix W describes the system's development (queue length $\hat{X}$ and environment $\hat{Y}$ process) if it is started empty, until the first customer enters the system.

The matrix $U^{(i,n)}$ describes the system's development from start of the ongoing service time of the, say, n-th admitted customer, until time τ_n^-; to be more precise, we describe an ongoing service and the subsequent departure but without the immediately following jump of the environment triggered by R.

We will use the following properties of the system and its describing process Z:

- the strong Markov (SM) property of Z,
- skip free to the left (SF) property of the system

$$P(Z(\zeta_1) = (n+j, m) | Z(0) = (n,k)) = 0 \qquad \forall j \geq 2, \tag{3.1.14}$$

- cadlag paths; in particular we are interested in the values of $Y(\tau_1^-)$, just before departure instants.

In the proofs of the following Proposition 3.1.8 and Proposition 3.1.9 we will often use distributive law for matrix addition and multiplication. It holds obviously for finite dimensional matrices, but for infinite dimensional matrices it is not always applicable and we need Lemma 3.1.7.

Lemma 3.1.7. *Let $M \in [0,1]^{K\times K}$ be arbitrary substochastic matrix on a countable state space K, $V := (v(k,m) : k, m \in K) \in \mathbb{R}^{K\times K}$ be an arbitrary generator and $a \in \mathbb{R}$ any constant. Then it holds*

$$(-\operatorname{diag}(V) + aI_W)M - (V - \operatorname{diag}(V))M = (aI_W - V)M. \tag{3.1.15}$$

Proof. Each (k,m)-th entry of the matrix on the left side of the equation (3.1.15) is an absolute convergent series

$$\left(-v(k,k) + a1_{[k\in K_W]}\right) M_{km} - \Big(\sum_{h\in K\setminus\{k\}} v(k,h)M_{hm}\Big).$$

To prove the absolute convergence we use the fact that all non-diagonal values of V are positive, that the diagonal elements of V are bounded, that the row sum of V is zero and that each entry of the stochastic matrix M is bounded by 1.

$$\begin{aligned}
&|\left(-v(k,k) + a1_{[k\in K_W]}\right) M_{km}| + \Big(\sum_{h\in K\setminus\{k\}} |v(k,h)M_{hm}|\Big) \\
\le &\underbrace{|\left(-v(k,k) + a\right)\cdot 1|}_{<\infty} + \underbrace{\Big(\sum_{h\in K\setminus\{k\}} |v(k,h)\cdot 1| \Big)}_{=v(k,k)<\infty}.
\end{aligned}$$

Because of the absolute convergence the difference on the left side of the equation (3.1.15) is well defined and we can use distributive law

$$(-\operatorname{diag}(V) + aI_W)M - (V - \operatorname{diag}(V))M = \left((-\operatorname{diag}(V) + aI_W) - (V - \operatorname{diag}(V))\right) M$$

to obtain the right side of the equation. □

Proposition 3.1.8. *For the matrix $W = (W_{km} : k, m \in K)$ from Lemma 3.1.6 it holds*

$$(\lambda I_W - V)W = \lambda I_W. \tag{3.1.16}$$

Proof. Recall that σ_1 denotes the arrival time of the first customer which is admitted to the system, which implies that at time σ_1 the environment is in a non-blocking state, and ζ_1 is the first jump time of the system which can be triggered only by V or by an arrival conditioned on $\hat{Y}$ being in K_W. It follows for $m \in K_W$

$$\begin{aligned}
W_{km} &= P(Z(\sigma_1) = (1,m)|Z(0) = (0,k)) \\
&= \sum_{h\in K\setminus\{k\}} P(Z(\sigma_1) = (1,m) \cap Z(\zeta_1) = (0,h)|Z(0) = (0,k)) \\
&\quad +\delta_{km}P(Z(\zeta_1) = (1,m)|Z(0) = (0,k)) \\
&= \sum_{h\in K\setminus\{k\}} P(Z(\sigma_1) = (1,m)|Z(\zeta_1) = (0,h), Z(0) = (0,k)) \\
&\quad \cdot P(Z(\zeta_1) = (0,h)|Z(0) = (0,k)) \\
&\quad +\delta_{km}P(Z(\zeta_1) = (1,m)|Z(0) = (0,k))
\end{aligned}$$

$$\overset{SM}{=} \sum_{h \in K\setminus\{k\}} \underbrace{P(Z(\sigma_1) = (1,m)|Z(0) = (0,h))}_{W_{hm}} \underbrace{P(Z(\zeta_1) = (0,h)|Z(0) = (0,k))}_{=\frac{v(k,h)}{-v(k,k)+\lambda 1_{[k\in K_W]}}}$$

$$+\delta_{km} \underbrace{P(Z(\zeta_1) = (1,m)|Z(0) = (0,k))}_{=\frac{\lambda}{-v(k,k)+\lambda 1_{[k\in K_W]}} 1_{[k\in K_W]}}$$

$$\iff W_{km} = \frac{1}{-v(k,k) + \lambda 1_{[k\in K_W]}} \Big(\sum_{h\in K} v_{kh}(1-\delta_{kh}) W_{hm} + \delta_{km}\lambda 1_{[k\in K_W]} \Big).$$

The equations above can be written in matrix form

$$W = (-\operatorname{diag}(V) + \lambda I_W)^{-1} \left((V - \operatorname{diag}(V))W + \lambda I_W \right)$$

$$\iff (-\operatorname{diag}(V) + \lambda I_W)W = (V - \operatorname{diag}(V))W + \lambda I_W$$

$$\iff (-\operatorname{diag}(V) + \lambda I_W)W - (V - \operatorname{diag}(V))W = \lambda I_W \tag{3.1.17}$$

$$\overset{\text{Lemma 3.1.7}}{\iff} (\lambda I_W - V)W = \lambda I_W. \tag{3.1.18}$$

□

Proposition 3.1.9. *For the matrices $U^{(i,n)} = (U^{(i,n)}_{km} : k, m \in K)$ from Lemma 3.1.6 it holds*

$$((\lambda + \mu(i))I_W - V)U^{(i,0)} = \mu(i) I_W \tag{3.1.19}$$

and

$$((\lambda + \mu(i))I_W - V)U^{(i,n+1)} = \lambda I_W U^{(i+1,n)}. \tag{3.1.20}$$

Proof. Note that τ_1 is the first departure time and ζ_1 is the first jump time of the system, and if this jump is triggered by a departure than $\zeta_1 = \tau_1$.

For $U^{(i,0)}$ with $i > 0$ it holds for $k \in K$ and $m \in K_W$:

$$\begin{aligned}
&U^{(i,0)}_{km} \\
&= P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i-1, m)|Z(0) = (i,k)\right) \\
&= \sum_{h\in K\setminus\{k\}} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i-1, m) \cap Z(\zeta_1) = (i,h)|Z(0) = (i,k)\right) \\
&\quad +\delta_{km} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i-1, k)|Z(0) = (i,k)\right) \\
&= \sum_{h\in K\setminus\{k\}} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i-1, m)|Z(\zeta_1) = (i,h), Z(0) = (i,k)\right) \\
&\quad \cdot P\left(Z(\zeta_1) = (i,h)|Z(0) = (i,k)\right) \\
&\quad +\delta_{km} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i-1, k)|Z(0) = (i,k)\right) \\
&\overset{SM}{=} \sum_{h\in K\setminus\{k\}} \underbrace{P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i-1, m)|Z(0) = (i,h)\right)}_{=U^{(i,0)}_{hm}} \frac{v(k,h)}{-v(k,k) + (\lambda + \mu(i))1_{[k\in K_W]}} \\
&\quad +\delta_{km} \frac{\mu(i)}{-v(k,k) + (\lambda + \mu(i))1_{[k\in K_W]}}
\end{aligned}$$

$$\begin{aligned}\iff\quad & U_{km}^{(i,0)}\\ &= \frac{1}{-v(k,k)+(\lambda+\mu(i))1_{[k\in K_W]}}\left(\sum_{h\in K} v(k,h)(1-\delta_{kh})U_{hm}^{(i,0)}+\delta_{km}\mu(i)1_{[k\in K_W]}\right).\end{aligned}$$

We write the equation above in a matrix form

$$U^{(i,0)} = \big(-diag(V)+(\lambda+\mu(i))I_W\big)^{-1}\cdot\Big((V-diag(V))U^{(i,0)}+\mu(i)I_W\Big)$$

$$\iff \big(-diag(V)+(\lambda+\mu(i))I_W\big)U^{(i,0)} = (V-diag(V))\,U^{(i,0)}+\mu(i)I_W \tag{3.1.21}$$

$$\overset{\text{Lemma 3.1.7}}{\iff} \big((\lambda+\mu(i))I_W - V\big)U^{(i,0)} = \mu(i)I_W. \tag{3.1.22}$$

Next we calculate for $n \geq 0$ and $1 \leq i$ the elements of the matrix $U_{km}^{(i,n+1)}$

$$\begin{aligned}
& U_{km}^{(i,n+1)}\\
&= P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n+1-1,m)|Z(0)=(i,k)\right)\\
&= \sum_{j=0}^{n+1}\sum_{h\in K} 1_{[(j,h)\neq(i,k)]}P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)\cap Z(\zeta_1)=(j,h)|Z(0)=(i,k)\right)\\
&\overset{SF}{=} \sum_{h\in K\backslash\{h\}} P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)\cap Z(\zeta_1)=(i,h)|Z(0)=(i,k)\right)\\
&\quad +P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)\cap Z(\zeta_1)=(i+1,k)|Z(0)=(i,k)\right)\\
&= \sum_{h\in K\backslash\{h\}} P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)|Z(\zeta_1)=(i,h),Z(0)=(i,k)\right)\\
&\quad \cdot P\left(Z(\zeta_1)=(i,h)|Z(0)=(i,k)\right)\\
&\quad +P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)|Z(\zeta_1)=(i+1,k),Z(0)=(i,k)\right)\\
&\quad \cdot P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+1,k)|Z(0)=(i,k)\right)\\
&\overset{SM}{=} \sum_{h\in K\backslash\{h\}} P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)|Z(0)=(i,h)\right)\\
&\quad \cdot P((Z(\zeta_1)=(i,h)|Z(0)=(i,k)))\\
&\quad +P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)|Z(\zeta_1)=(i+1,k)\right)\\
&\quad \cdot P\left(Z(\zeta_1)=(i+1,k)|Z(0)=(i,k)\right)\\
&= \sum_{h\in K\backslash\{h\}} \underbrace{P(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n,m)|Z(0)=(i,h))}_{=U_{hm}^{(i,n+1)}}\\
&\quad \cdot\frac{v(k,h)}{-v(k,k)+(\mu(i)+\lambda)1_{[k\in K_W]}}\\
&\quad +\underbrace{P\left(\left(X(\tau_1),Y(\tau_1^-)\right)=(i+n+1-1,m)|Z(0)=(i+1,k)\right)}_{=U_{km}^{(i+1,n)}}\\
&\quad \cdot 1_{[k\in K_W]}\frac{\lambda}{-v(k,k)+(\mu(i)+\lambda)1_{[k\in K_W]}}
\end{aligned}$$

$$U_{km}^{(i,n+1)} = \frac{1}{-v(k,k) + (\mu(i)+\lambda)1_{[k\in K_W]}} \cdot \left(\sum_{h\in K} v(k,h)(1-\delta_{kh})U_{hm}^{(i,n+1)} + \lambda 1_{[k\in K_W]} U_{km}^{(i+1,n)}\right).$$

The last equation can be written in matrix form as

$$U^{(i,n+1)} = (-\operatorname{diag}(V) + (\lambda+\mu(i))I_W)^{-1} \cdot \Big((V - \operatorname{diag}(V))\, U^{(i,n+1)} + \lambda I_W U^{(i+1,n)}\Big)$$

$$\iff (-\operatorname{diag}(V) + (\lambda+\mu(i))I_W)\, U^{(i,n+1)} = (V - \operatorname{diag}(V))\, U^{(i,n+1)} + \lambda I_W U^{(i+1,n)} \tag{3.1.23}$$

$$\overset{\text{Lemma 3.1.7}}{\iff} ((\lambda+\mu(i))I_W - V)U^{(i,n+1)} = \lambda I_W U^{(i+1,n)}. \tag{3.1.24}$$

□

Corollary 3.1.10. *If the matrices $(\lambda I_W - V)^{-1}$ are invertible for any positive λ then it holds*

$$W = \lambda(\lambda I_W - V)^{-1} I_W. \tag{3.1.25}$$

Proof. (3.1.25) follows directly from (3.1.16). □

We are now ready to evaluate the steady-state equations (3.1.8) of $\hat{Z}$. Because we have a Poisson-λ arrival stream, the marginal steady state (2.1.8) of the continuous time queue length process X is

$$\xi = \left(\xi(n) := C^{-1}\prod_{i=1}^{n} \frac{\lambda}{\mu(i)} : n \in \mathbb{N}_0\right).$$

Recall (3.1.8)

$$\hat{\pi}^{(0)} B^{(n)} + \sum_{i=1}^{n+1} \hat{\pi}^{(i)} A^{(i,n-i+1)} = \hat{\pi}^{(n)}, \qquad n \in \mathbb{N}_0,$$

and the decomposition from Lemma 3.1.6:

$$A^{(i,n)} = U^{(i,n)} R \quad \text{and} \quad B^{(n)} = W U^{(1,n)} R.$$

The *conjectured product form steady state* will eventually be realized as

$$\hat{\pi}(n,k) = \xi(n)\cdot\hat{\theta}(k), \text{ for } (n,k)\in E, \text{ and } \hat{\pi}^n = \xi(n)\cdot\hat{\theta}, \text{ for } n \in \mathbb{N}_0,$$

with $\hat{\theta}(k) = 0$ for some $k \in K$.

The *idea of the proof* is: The steady-state equation is transformed into

$$\xi(n)\cdot\hat{\theta} = \xi(0)\cdot\hat{\theta}\cdot W\cdot U^{(1,n)}\cdot R + \sum_{i=1}^{n+1} \xi(n)\cdot\hat{\theta}\cdot U^{(i,n-i+1)}\cdot R, \qquad \forall n \in \mathbb{N}_0. \tag{3.1.26}$$

We insert $\xi(n)$, cancel C^{-1}, and obtain the *"environment equations"*

$$\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)\cdot\hat{\theta}=\hat{\theta}\cdot W\cdot U^{(1,n)}\cdot R+\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\hat{\theta}\cdot U^{(i,n-i+1)}\cdot R \qquad \forall n\in\mathbb{N}_0, \tag{3.1.27}$$

which we may consider as a sequence of equations with vector of unknowns $\hat{\theta}$. Then, using properties of the matrices W and $U^{(i,n)}$ we show that the proposed $\hat{\theta}=(\theta I_W\mathbf{e})^{-1}\cdot\theta I_W R$ solves these environment equations.

Lemma 3.1.11. *Let θ be a positive stochastic solution of the equation $\theta\left(\lambda I_W\left(R-I\right)+V\right)=0$, where λ, I_W, R and V are parameters of an ergodic loss system in continuous time. Then the vector*

$$y:=(\theta I_W\mathbf{e})^{-1}\cdot\theta I_W R \tag{3.1.28}$$

is stochastic and it holds

$$y=(\theta I_W\mathbf{e})^{-1}\cdot\frac{1}{\lambda}\theta\left(\lambda I_W-V\right). \tag{3.1.29}$$

Proof. The vector $y=(\theta I_W\mathbf{e})^{-1}\cdot\theta I_W R$ is stochastic because θ is stochastic, $I_W R$ is sub-stochastic, and $(\theta I_W\mathbf{e})=(\theta I_W\underbrace{R\mathbf{e}}_{=\mathbf{e}})$ is the appropriate positive normalization constant.

For finite set K the property (3.1.29) follows from the obvious transformations

$$\theta\left(\lambda I_W\left(R-I\right)+V\right)=0\iff\theta\lambda I_W R-\theta\lambda I_W+\theta V=0\iff\lambda\theta I_W R=\theta\left(\lambda I_W-V\right) \tag{3.1.30}$$

and multiplying both sides with $(\theta I_W\mathbf{e})^{-1}\frac{1}{\lambda}$. But when K is infinite, we need to prove that we can use distributive law for θ:

Each k-th vector element of vectors $\theta\left(\lambda I_W\left(R-I\right)+V\right)$, $\theta\lambda I_W R-\theta\lambda I_W+\theta V$, $\lambda\theta I_W R$, and $\theta\left(\lambda I_W-V\right)$ in (3.1.30) is absolute convergent series, whose sum of absolute terms is bounded by

$$\sum_{h\in K}|\lambda\theta(h)R(h,k)|\,1_{[k\in K_W]}+|\theta(k)\lambda 1_{[k\in K_W]}|+\sum_{h\in K}|\theta(h)v(h,k)|,$$

which is bounded too:

$$\begin{aligned}&\sum_{h\in K}|\lambda\theta(h)R(h,k)|\,1_{[h\in K_W]}+\theta(k)\lambda 1_{[k\in K_W]}+\sum_{h\in K}|\theta(h)v(h,k)|\\
&=\lambda\sum_{h\in K}\left(\theta(h)(R(h,k)-\delta_{hk}\right)1_{[h\in K_W]}+2\theta(k)\lambda 1_{[k\in K_W]}+\sum_{h\in K}\theta(h)v(h,k)-2\theta(k)v(k,k)\\
&=\underbrace{\lambda\sum_{h\in K}\theta(h)(R(h,k)-\delta_{hk})1_{[h\in K_W]}+\sum_{h\in K}\theta(h)v(h,k)}_{=0}+\underbrace{2\lambda\theta(k)1_{[k\in K_W]}-2\theta(k)v(k,k)}_{<\infty}.\end{aligned}$$

Therefore the transformation (3.1.30) is valid also for infinite K. □

Lemma 3.1.12. *Let θ be a positive stochastic solution of the equation $\theta\left(\lambda I_W(R-I)+V\right)=0$, where λ, I_W, R and V are parameters of an ergodic loss system in continuous time. Then $y=(\theta I_W\mathbf{e})^{-1}\cdot\theta I_W R$ solves the environment equations* (3.1.27)*, that is*

$$\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)\cdot y=y\cdot W\cdot U^{(1,n)}\cdot R+\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot y\cdot U^{(i,n-i+1)}\cdot R,\qquad\forall n\in\mathbb{N}_0.\quad(3.1.31)$$

Proof. From Lemma 3.1.11 follows that y is a stochastic vector. Vector y and every matrix on the right side of the equation (3.1.31) contain only non-negative elements, additionally all series on the right side are bounded by the finite values $\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)y(k)$ on the left side, therefore the following transformation is the same for finite and infinite K

$$\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)\cdot y=\left(y\cdot W\cdot U^{(1,n)}+\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot y\cdot U^{(i,n-i+1)}\right)\cdot R.\qquad(3.1.32)$$

From Lemma 3.1.11 follows $y=(\theta I_W\mathbf{e})^{-1}\theta\frac{1}{\lambda}\left(\lambda I_W-V\right)$. We insert $y=(\theta I_W\mathbf{e})^{-1}\cdot\theta\frac{1}{\lambda}\left(\lambda I_W-V\right)$ into the right side of equation (3.1.32), insert $y=(\theta I_W\mathbf{e})^{-1}\cdot\theta I_W R$ into the left side of the equation (3.1.32) and multiply both sides with $(\theta I_W\mathbf{e})\lambda$.

$$\begin{aligned}&\lambda\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)\cdot\theta I_W R\\&=\left(\theta\left(\lambda I_W-V\right)\cdot W\cdot U^{(1,n)}+\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}\right)\cdot R.\end{aligned}$$

Using the property (3.1.16) of the matrix W we get

$$\lambda\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)\cdot\theta I_W R=\left[\theta\lambda I_W\cdot U^{(1,n)}+\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}\right]\cdot R.\qquad(3.1.33)$$

Now we iteratively reduce the number of the summands in the squared brackets for $n\geq 1$

$$\begin{aligned}&\theta\lambda I_W\cdot U^{(1,n)}+\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}\\&=\theta\lambda I_W\cdot U^{(1,n)}+\frac{\lambda}{\mu(1)}\cdot\theta\left(\lambda I_W-V\right)U^{(1,n)}+\sum_{i=2}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}\\&=\frac{\lambda}{\mu(1)}\Big(\underbrace{\theta\mu(1)I_W+\theta\left(\lambda I_W-V\right)}_{(*)}\Big)U^{(1,n)}+\sum_{i=2}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}.\end{aligned}$$

We can apply distributive law for θ in (*) because every k-th element of $\theta\mu(1)I_W$ and $\theta\lambda I_W$ is bounded and in the proof of Lemma 3.1.11 we showed that every k-th element of θV is a an absolute convergent series

$$\frac{\lambda}{\mu(1)}\theta\left(\mu(1)I_W+\lambda I_W-V\right)U^{(1,n)}+\sum_{i=2}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}.$$

Using the property (3.1.20) of the matrix $U^{(1,n)}$we get

$$\frac{\lambda}{\mu(1)}\theta\lambda I_W U^{(2,n-1)}+\sum_{i=2}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}.$$

Using same technique we remove the next summand for $n\geq 2$

$$\begin{aligned}
&\frac{\lambda}{\mu(1)}\theta\lambda I_W U^{(2,n-1)}+\left(\prod_{j=1}^{2}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(2,n-2+1)}\\
&+\sum_{i=3}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}\\
=&\left(\prod_{j=1}^{2}\frac{\lambda}{\mu(i)}\right)\theta\underbrace{\left(\mu(2)I_W+\lambda I_W-V\right)U^{(2,n-1)}}_{=\lambda I_W U^{(3,n-2)}}\\
&+\sum_{i=3}^{n+1}\left(\prod_{j=1}^{i}\frac{\lambda}{\mu(i)}\right)\cdot\theta\left(\lambda I_W-V\right)U^{(i,n-i+1)}.
\end{aligned}$$

After applying this technique $n-2$ more times we arrive at

$$\begin{aligned}
&\left(\prod_{j=1}^{n}\frac{\lambda}{\mu(i)}\right)\theta\lambda I_W U^{(n+1,0)}+\prod_{j=1}^{n+1}\frac{\lambda}{\mu(i)}\theta\left(\lambda I_W-V\right)U^{(n+1,0)}\\
=&\left(\prod_{j=1}^{n+1}\frac{\lambda}{\mu(i)}\right)\theta\left(\mu(n+1)I_W+\lambda I_W-V\right)U^{(n+1,0)}
\end{aligned}$$

and the equation (3.1.33) can be written as

$$\begin{aligned}
\lambda\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)\cdot\theta I_W R=&\left(\prod_{j=1}^{n+1}\frac{\lambda}{\mu(i)}\right)\theta\left(\left(\mu(n+1)I_W+\lambda I_W-V\right)U^{(n+1,0)}\right)R\\
\overset{(3.1.19)}{\Longleftrightarrow}\ &\lambda\theta I_W R=\theta\frac{\lambda}{\mu(n+1)}\mu(n+1)I_W R.
\end{aligned}$$

□

Lemma 3.1.13. *Let θ be a positive stochastic solution of the equation $\theta(\lambda I_W(I-R)+V)=0$, $y := (\theta I_W \mathbf{e})^{-1} \cdot \theta I_W R$, where λ, I_W, R and V are parameters of an ergodic loss system in continuous time and W is the matrix from Lemma 3.1.6. Then*

$$\theta I_W = (\theta I_W \mathbf{e}) y W \tag{3.1.34}$$

and y is a stochastic solution of the equation

$$yWR = y. \tag{3.1.35}$$

If $(\lambda I_W - V)$ is invertible then

$$\theta = \left(y \left(I_W - \frac{1}{\lambda} V \right)^{-1} \mathbf{e} \right)^{-1} y \left(I_W - \frac{1}{\lambda} V \right)^{-1}. \tag{3.1.36}$$

Proof. From definition of y and (3.1.29) in Lemma 3.1.11 follows

$$\frac{1}{\lambda} \theta \left(\lambda I_W - V \right) = (\theta I_W \mathbf{e}) y. \tag{3.1.37}$$

To obtain (3.1.34) we multiply both sides of (3.1.37) with W from the right and use property (3.1.16) of the matrix W

$$\frac{1}{\lambda} \theta \underbrace{\left(\lambda I_W - V \right) W}_{=\lambda I_W} = (\theta I_W \mathbf{e}) y W.$$

For (3.1.35) we multiply both sides of the equation (3.1.34) with $(\theta I_W \mathbf{e})^{-1}$ from the left and R from the right

$$(\theta I_W \mathbf{e})^{-1} \theta I_W R = yWR \Longleftrightarrow y = yWR.$$

By construction, y is stochastic.

The matrix $(\lambda I_W - V)$ is invertible if and only if $\left(I_W - \frac{1}{\lambda}V\right)^{-1}$ is invertible. The equation (3.1.36) follows if we multiply both side of equation (3.1.37) with $(\lambda I_W - V)^{-1}$ from the right

$$\theta = (\theta I_W \mathbf{e}) y \left(I_W - \frac{1}{\lambda} V \right)^{-1}.$$

We know that θ is stochastic. Therefore we can replace the normalization constant $(\theta I_W \mathbf{e})$ by $\left(y \left(I_W - \frac{1}{\lambda} V \right)^{-1} \mathbf{e} \right)^{-1}$. □

Proposition 3.1.14. *Consider the ergodic Markov process $Z = (Z(t) : t \geq 0)$ which describes the $M/M/1/\infty$ loss system in a random environment. Let $L := \{k \in K : \exists\ m \in K_W : R(m,k) > 0\}$ the set of states of the environment which can be reached from K_W by a one-step jump governed by R. Then the states in $\mathbb{N}_0 \times (K \setminus L)$ are inessential for $\hat{Z}$ and consequently for all $n \in \mathbb{N}_0$*

$$\hat{\pi}(n,k) = 0, \qquad \forall k \in (K \setminus L). \tag{3.1.38}$$

The steady-state distribution $\hat{\pi}$ is unique.

Proof. We realize from the dynamics of the system that $\mathbb{N}_0 \times L$ are the only states that can be entered just after a departure instant. So, if $\hat{Z}$ is started in some state $\mathbb{N}_0 \times (K \setminus L)$ these states states will never be visited again by $\hat{Z}$ and are therefore inessential. And we have (3.1.38).

The uniqueness follows from (3.1.38) and irreducibility of $\mathbf{P}$ on the class $\mathbb{N}_0 \times L$. We prove irreducibility of $\mathbf{P}$ on the class $\mathbb{N}_0 \times L$ as follows: Let (n,k) and (n',m) be any states with $n,n' \in \mathbb{N}_0$ and $k,m \in L$ then there exists $m' \in K$ such that $R(m',m) > 0$. Because Z is ergodic, there exist a finite sequence of states $\Big((n_i,k_i)\Big)_{i\in\{0,\dots,s\}}$ from (n,k) to $(n'+1,m')$ with

$$(n_0,k_0) = (n,k) \text{ and } (n_s,k_s) = (n'+1,m')$$

such that

$$q((n_i,k_i),(n_{i+1},k_{i+1})) > 0 \qquad \forall i \in \{0,\dots,s-1\}.$$

The sequence contains finite number of transitions downwards $t \in \mathbb{N}_0$, i.e. from some level $\ell+1$ to level ℓ. When finally, the continuous process reaches the state $(n'+1,m')$ it jumps with positive probability $\frac{1}{v(m',m')+\lambda+\mu(n'+1)} R(m',k)$ to state (n,k). Therefore $(\mathbf{P}^{t+1})_{((n,k),(n',m))} > 0$. $\square$

Remark 3.1.15. In case of $K \setminus L \neq \emptyset$ $\hat{Z}$ is not irreducible on E, hence not ergodic, although Z is ergodic on E. Furthermore, in general $\hat{Z}$ is even on the reduced state space $\hat{E} := \mathbb{N}_0 \times L$ not ergodic. The reason is, that $\hat{Z}$ may have periodic classes as the Example 3.1.18 shows.

The result (3.1.38) in Proposition 3.1.14 is rather general and also hold for $M/G/1/\infty$ loss systems as described in Section 3.2. The proof carries over directly.

Theorem 3.1.16. *Consider the ergodic Markov process $Z = (Z(t) : t \geq 0)$ which describes the $M/M/1/\infty$ loss system in a random environment with steady-state probability π from* (2.1.4) *with $\pi(n,k) = \xi(n)\theta(k)$, $(n,k) \in E$.*

(a) The Markov chain $\hat{Z} = (\hat{Z}(n) : n \in \mathbb{N}_0)$ embedded at departure instants of Z has the stationary distribution $\hat{\pi}$ of product form

$$\hat{\pi}(n,k) = \xi(n)\hat{\theta}(k), \qquad (n,k) \in E. \tag{3.1.39}$$

Here $\xi = (\xi(n) : n \in \mathbb{N}_0)$ is the probability distribution

$$\xi(n) := C^{-1}\left(\prod_{i=1}^{n} \frac{\lambda}{\mu(i)}\right), \qquad n \in \mathbb{N}_0, \tag{3.1.40}$$

with normalization constant C^{-1} and

$$\hat{\theta} = (\theta I_W \mathbf{e})^{-1} \cdot \theta I_W R \tag{3.1.41}$$

which is independent of the values $\mu(n)$.

(b) $\hat{\theta}$ is a stochastic solution of the equation

$$\hat{\theta} W R = \hat{\theta}, \tag{3.1.42}$$

(c) If the matrix $(\lambda I_W - V)$ is invertible, for the marginal stationary distribution θ of Y in continuous time and the marginal stationary distribution $\hat{\theta}$ of $\hat{Y}$ of the embedded Markov chain $\hat{Z}$ it holds

$$\theta = \left(\hat{\theta}\left(I_W - \frac{1}{\lambda}V\right)^{-1}\mathbf{e}\right)^{-1}\hat{\theta}\left(I_W - \frac{1}{\lambda}V\right)^{-1}. \tag{3.1.43}$$

Proof. Part (a). We show that the product form distribution (3.1.39) with marginal distributions (3.1.40) and the solution $\hat{\theta}$ of (3.1.42) solves the steady-state equations (3.1.8).

$$\begin{aligned}
&\hat{\pi}^{(n)} = \hat{\pi}^{(0)}B^{(1,n)} + \sum_{i=1}^{n+1}\hat{\pi}^{(n)}A^{(i,n-i+1)}\\
\Longleftrightarrow\ & \xi(n)\hat{\theta} = \xi(0)\hat{\theta}WU^{(1,n)}R + \sum_{i=1}^{n+1}\xi(n)\hat{\theta}U^{(i,n-i+1)}R\\
\Longleftrightarrow\ & \frac{\xi(n)}{\xi(0)}\hat{\theta} = \hat{\theta}WU^{(1,n)}R + \sum_{i=1}^{n+1}\frac{\xi(n)}{\xi(0)}\hat{\theta}U^{(i,n-i+1)}R.
\end{aligned}$$

The last equation is (3.1.31) in Lemma 3.1.12 with $y = \hat{\theta}$ and $\xi(n) = C^{-1}\left(\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\right)$ and we proved that $\hat{\pi}(n,k) = \xi(n)\hat{\pi}(k)$ is a solution of the steady-state equation (3.1.5).The uniqueness of $\hat{\pi}$ is given in Proposition 3.1.14.

Part (b) and (c). We have proven in part (a) that $\hat{\theta} = (\theta I_W\mathbf{e})^{-1}\cdot\theta I_W R$, so we can apply Lemma 3.1.13 and get both (3.1.42) and (3.1.43). □

Corollary 3.1.17. *Consider the ergodic Markov process $Z = (Z(t) : t \geq 0)$ which describes an $M/M/1/\infty$ loss system in a random environment with steady-state probability π from* (2.1.4) *with finite environment space K. Then the stochastic solution of the equation* (3.1.42) *is unique and it is the marginal steady-state distribution $\hat{\theta}$ of the environment. The equation* (3.1.42) *can be written in the form*

$$\hat{\theta}\underbrace{\lambda(\lambda I_W - V)^{-1}I_W}_{=W \textit{ from } (3.1.25)}R = \hat{\theta}. \tag{3.1.44}$$

Proof. Because the environment state space K is finite, according to Proposition A.2.2 the matrix $(\lambda I_W - V)$ is invertible. Using (3.1.25) in Corollary 3.1.10 the equation (3.1.42) can be written as (3.1.44).

According to Theorem 3.1.16 the marginal distribution $\hat{\theta}$ solves the equation (3.1.42). To prove the corollary we show that the equation (3.1.42) has a unique stochastic solution.

Assume that $\hat{\theta}_1$ and $\hat{\theta}_2$ are different non-zero solutions of the equation (3.1.42), which is equivalent to

$$\begin{aligned}
&\hat{\theta}_i\underbrace{\lambda(\lambda I_W - V)^{-1}I_W}_{=W \text{ from } (3.1.25)}R = \hat{\theta}_i\underbrace{(\lambda I_W - V)^{-1}(\lambda I_W - V)}_{=I}\\
\Longleftrightarrow\ & \hat{\theta}_i(\lambda I_W - V)^{-1}\left(\lambda I_W(R - I) + V\right) = 0.
\end{aligned}$$

We define $x_1 := \hat{\theta}_1(\lambda I_W - V)^{-1}$ and $x_2 := \hat{\theta}_2(\lambda I_W - V)^{-1}$. Both x_1 and x_2 are solution of the continuous time steady-state equation $x_i\,(\lambda I_W(R-I)+V)=0$. Recall, that due to ergodicity of the process (X,Y), according to Corollary 2.1.10, the equation $\theta\,(\lambda I_W\,(R-I)+V)=0$ has a unique stochastic solution θ. This means that there exists some constant c such that $x_1 = cx_2$ holds. By definition of x_i it is equivalent to

$$x_1 - cx_2 = 0 \Longleftrightarrow (\hat{\theta}_1 - c\hat{\theta}_2)(\lambda I_W - V)^{-1} = 0.$$

Because $(\lambda I_W - V)^{-1}$ is bijective on the finite dimensional state space $\mathbb{R}^K$, it follows $\hat{\theta}_1 - c\hat{\theta}_2 = 0$ and thus the uniqueness of stochastic solution $\hat{\theta}$ of (3.1.42). □

Example 3.1.18. [SSD+06](See also Section 2.2.1.) We consider an M/M/1/∞-system with attached inventory, i.e. a single server with infinite waiting room under FCFS regime and an attached inventory under (r,S)-policy, which is set in this example to $r=0$.

There is a Poisson-λ-arrival stream, $\lambda > 0$. Customers request for an amount of service time which is exponentially distributed with mean $\mu > 0$.

The server needs for each customer exactly one item from the inventory. The on-hand inventory decreases by one at the moment of service completion. If the inventory is decreased to the reorder point $r=0$ after the service of a customer is completed, a replenishment order is instantaneously triggered. The replenishment lead times are i.i.d. exponentially distributed with parameter $\nu > 0$. The replenishment fills the inventory up to maximal inventory size $S > 0$.

During the time the inventory is depleted and the server waits for a replenishment order to arrive, no customers are admitted to join the queue ("lost sales").
All service, interarrival and lead times are assumed to be independent.
$X(t)$ is the number of customers present at the server at time $t \geq 0$, and $Y(t)$ is the on-hand inventory at time $t \geq 0$.

The state space of (X,Y) is $E = \{(n,k) : n \in \mathbb{N}_0, k \in K\}$, with $K = \{S, S-1, \ldots, 1, 0\}$, where $S < \infty$ is the maximal size of the inventory at hand.

The inventory management process under $(0,S)$-policy fits into the definition of the environment process by setting

$$K = \{S, S-1, \ldots, 1, 0\}, \qquad K_B = \{0\},$$

$$R(0,0) = 1, \quad R(k,k-1) = 1, \quad 1 \leq k \leq S\,, \qquad v(k,m) = \begin{cases} \nu, & \text{if } k=0, m=S, \\ 0, & \text{otherwise for } k \neq m. \end{cases}$$

The queueing-inventory process $Z = (X,Y)$ in continuous time is ergodic iff $\lambda < \mu$. The steady-state distribution $\pi = (\pi(n,k) : (n,k) \in E)$ of (X,Y) has product form

$$\pi(n,k) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\nu}\right)^n \theta(k),$$

where $\theta = (\theta(k) : k \in K)$ with normalization constant C is

$$\theta(k) \quad = \begin{cases} C^{-1}(\frac{\lambda}{\nu}), & k = 0, \\ C^{-1}(\frac{\lambda+\nu}{\lambda})^{k-1}, & k = 1, \ldots, r, \\ C^{-1}(\frac{\lambda+\nu}{\lambda})^{r}, & k = r+1, \ldots, S. \end{cases} \tag{3.1.45}$$

For the Markov chain $\hat{Z}$ embedded in Z at departure instants we have $L = \{0, 1, \ldots, S-1\}$ and therefore the states $\mathbb{N}_0 \times \{S\}$ are inessential.

From the dynamics of the system determined by the inventory management follows directly that $\hat{Z}$ is periodic with period S and that $\mathbb{N}_0 \times L$ is an irreducible closed set (the single essential class), which is positive recurrent iff $\lambda < \mu$ holds. $\mathbb{N}_0 \times L$ is partitioned into S subclasses $\mathbb{N}_0 \times \{k\}$ which are periodically visited

$$\ldots \to \mathbb{N}_0 \times \{S-1\} \to \mathbb{N}_0 \times \{S-2\} \to \ldots \to \mathbb{N}_0 \times \{0\} \to \mathbb{N}_0 \times \{S-1\} \ldots$$

The following corollary and examples demonstrate the versatility of the class of models under consideration and consequences for the interplay of θ for the continuous time setting and $\hat{\theta}$ for the embedded Markov chain due to special settings of the environment.

Corollary 3.1.19. *Consider an ergodic $M/M/1/\infty$ loss system in a random environment with any λ, $\mu(n)$, V, and R as defined in Section 2.1.1.*

(a) If $R = I$, then the conditional distribution $\hat{\theta}$ of θ conditioned on L,

$$\hat{\theta}(k) = \begin{cases} \frac{\theta(k)}{\theta(L)} & \text{if } k \in L, \\ 0 & \text{if } k \in K \setminus L, \end{cases} \qquad \text{with} \quad \theta(L) := \sum_{m \in L} \theta(m),$$

solves (3.1.42). *This shows that the embedded chain in this case reveals only the behaviour of the environment on L, i.e. we loose information incorporated in the continuous time description of the process.*

(b) If $V = 0$ then the set K_B of blocking states is empty, and therefore $I_W = I$ holds. Furthermore, R is irreducible and positive recurrent.
The marginal steady-state distribution θ of Y in continuous time is the stationary distribution of R, i.e., the solution of $\theta R = \theta$.
And finally it holds $\theta = \hat{\theta}$, i.e., θ solves on K (3.1.42), *which shows that the embedded chain exploits in this case the full information about the possible environment of the system.*

Proof. **(a)** Substituting $R = I$ into (3.1.41) leads to $\hat{\theta}(k) = (\theta I_W \mathbf{e})^{-1} \cdot \theta(k)$ and to $\hat{\theta}(k) > 0 \iff k \in K_W$. From $R = I$ we have $L = K_W$ and therefore $(\theta I_W \mathbf{e}) = \sum_{m \in K_W} \theta(m) = \theta(L)$.

(b) If $V = 0$ and $K_B \neq \emptyset$, then from ergodicity the environment process Y must enter K_B in finite time, but once the system entered a blocking state k it can never leave this because of $v(k, m) = 0$ for all $m \in K$. Furthermore, from ergodicity of Z with a similar argument, R must be irreducible and positive recurrent.

(2.1.18) then reduces to $\theta(\lambda(R - I)) = 0$ which is the steady-state equation for R.

We substitute $I_W = I$ and $V = 0$ into (3.1.16) and obtain $W = I$. For $W = I$ the equation (3.1.42) reduces to $\hat{\theta} R = \hat{\theta}$, which from irreducibility and positive recurrence has a unique stochastic solution $\hat{\theta}$. □

Part (b) of Theorem 3.1.16 demonstrates the link between marginal distributions θ and $\hat{\theta}$ and rises the question "How different can θ and $\hat{\theta}$ be?" The following theorem proves that they can be arbitrarily different.

Theorem 3.1.20. *Let θ be any positive distribution on a finite set K, $K_W \subset K$ any non-empty subset of K, and $\hat{\theta}$ any distribution on K. Then there exists an ergodic loss system with environment space $K = K_W \uplus K_B$ and steady-state distribution $\pi(n,k) = \xi(n)\theta(k)$ of the continuous time process, whereas $\hat{\pi}(n,k) = \xi(n)\hat{\theta}(k)$ is the steady-state distribution of the associated Markov chain embedded at departure epochs.*

Proof. We will construct an ergodic system with marginal steady-state distributions θ in continuous time and marginal steady-state distribution $\hat{\theta}$ of the embedded Markov chain for the environment. According to Theorem 3.1.16 (a) for this system the equation (3.1.41) must hold. With normalization constant $c := \theta I_W \mathbf{e}$ this is

$$\hat{\theta} = (\theta I_W \mathbf{e})^{-1} \cdot \theta I_W R = c^{-1} \theta I_W R. \tag{3.1.46}$$

First we construct R with identical rows equal to $\hat{\theta}$

$$R := \begin{pmatrix} \hat{\theta} \\ \vdots \\ \hat{\theta} \end{pmatrix}.$$

For any stochastic vector x holds $\hat{\theta} = xR$, in particular (3.1.46) is true for the stochastic vector $c^{-1}\theta I_W$.

Now we construct V, λ and μ. Let $\lambda \in \mathbb{R}$ be any positive number with

$$\lambda < \begin{cases} \frac{\min_{k\in K} \theta(k)}{\max_{k\in K} |\theta I_W(k) - c\hat{\theta}(k)|} & \text{if } \theta I_W \neq c\hat{\theta}, \\ \infty & \text{if } \theta I_W = c\hat{\theta}, \end{cases}$$

and $\mu \in \mathbb{R}$ be any number greater than λ.

To construct the generator matrix V, we define $w := \lambda(\theta I_W - c\hat{\theta}) + \theta$. The vector w is positive

$$\begin{aligned} w(k) \geq & -\lambda|\theta I_W(k) - c\hat{\theta}(k)| + \theta(k) \\ & \begin{cases} > -\min_{k\in K}(\theta(k)) + \theta(k) \geq 0 & \text{if } (\theta I_W(k) - \hat{\theta}(k)) \neq 0, \\ = \theta(k) > 0 & \text{if } (\theta I_W(k) - \hat{\theta}(k)) = 0, \end{cases} \end{aligned}$$

and stochastic because of

$$w\mathbf{e} = (\lambda(\theta I_W - c\hat{\theta}) + \theta)\mathbf{e} = \lambda \underbrace{\theta I_W \mathbf{e}}_{=c} - \lambda c \underbrace{\hat{\theta}\mathbf{e}}_{=1} + \underbrace{\theta\mathbf{e}}_{=1} = 1.$$

Finally we define the stochastic matrix $\mathcal{W} := \begin{pmatrix} w \\ \vdots \\ w \end{pmatrix}$ and the generator matrix $V := \mathcal{W} - I$. For any stochastic vector x holds

$$xV = x(\mathcal{W} - I) = w - x = \lambda(\theta I_W - c\hat{\theta}) + \theta - x,$$

and so θ is the solution of (2.1.18), which is $\theta(\lambda I_W(R - I) + V) = 0$

$$\iff \lambda \underbrace{\theta I_W R}_{=c\hat{\theta}} - \lambda\theta I_W + \underbrace{\theta V}_{=\lambda(\theta I_W - c\hat{\theta}) + \theta - \theta} = 0.$$

We have $\lambda < \mu$, so (2.1.5) holds. Furthermore, due to construction of $\mathcal{W}$ all non-diagonal matrix elements of $Q_{\text{red}} = \lambda I_W(R-I)+V$ are positive, hence Q_{red} is irreducible and a positive recurrent generator of some Markov process with unique stochastic solution of $\theta Q_{\text{red}} = 0$. So we have proven the theorem. $\square$

Remark 3.1.21. When $K_W = K$ the process constructed in Theorem 3.1.20 is a birth-death process in random environment with separable steady-state distribution. It has Poisson input stream and exponentially distributed service times without interruptions.

3.2. $M/G/1/\infty$ queueing system in a random environment

Vineetha [Vin08] extended the theory of integrated queueing-inventory models with exponential service times as described in [SSD+06] to systems with i.i.d. service times which follow a general distribution. The lead time is exponential and during stock-out periods lost sales occur. Her approach was classical in that she considered the continuous time Markovian state process at departure instants of customers.

In this section we revisit some of Vineetha's [Vin08] models. We prove some of our results for queues with general environments from the previous sections on $M/M/1/\infty$ systems in the $M/G/1/\infty$ framework, which includes a partial extensions of Vineetha's queueing-inventory systems to queues with state dependent service speeds and with non-exponential service times.

Our framework is as in Section 3.1: Consider the system at departure instants and utilize Markov chain analysis.

Our main aim is to identify conditions which enforce the systems to stabilize in a way that the queue and the environment decouple in the sense that the stationary queue length and environment of the embedded Markov Chain behave independently for a fixed time instant, i.e., a product form equilibrium exists.

It will come out that this is not always possible, but we are able to provide sufficient conditions for the existence of product form equilibria.

3.2.1. $M/G/1/\infty$ queueing systems with state dependent service intensities

We first describe a pure queueing model in continuous time which is of $M/G/1/\infty$ type, under FCFS regime, where the single server works with different queue length dependent speeds ("service intensities"), and the customers' service requests are queue length dependent as well.

A review of $M/G/1/\infty$ queueing systems with state dependent arrival and service intensities, which are related to the model described here, and their asymptotic and equilibrium behaviour is provided in the survey of Dshalalow [Dsh97].

The arrival stream is Poisson-λ. When a customer enters the single server seeing $n-1 \geq 0$ customers behind him, i.e., the queue length is n, his amount of requested service time is drawn according to a distribution function $G_n : [0,\infty) \to [0,1]$ with $G_n(0) = 0$. The set of all interarrival times and service time requests is an independent collection of variables.

The server works with queue length dependent service speeds $c(n) > 0$, i.e., when at time $t \geq 0$ there are $X(t) = n > 0$ customers in the system (n including the one in service), and if the residual service request of the customer in service at time t is $R_{\mathrm{S}}(t) = r > 0$, then at time $t + \varepsilon$ his residual service request is

$$R_{\mathrm{S}}(t + \varepsilon) = r - \varepsilon \cdot c(n), \quad \text{if this is } > 0\,,$$

otherwise at time $t + \varepsilon$ his service expired and he has already departed from the system. For $X(t) = 0$ we define $R_{\mathrm{S}}(t) = 0$.

It is a standard observation that the process

$$(X, R_{\mathrm{S}}) = ((X(t), R_{\mathrm{S}}(t)) : t \geq 0)$$

is a homogeneous strong Markov process on state space $\mathbb{N}_0 \times \mathbb{R}_0^+$ (with cadlag paths).

With $\tau_0 = 0$ we will denote as in the previous sections by $\tau = (\tau_0, \tau_1, \dots)$ the sequence of departure times of customers. From process definition we conclude

$$R(\tau_n^-) = 0, \qquad \forall n \geq 1.$$

For $\tau_0 = 0$ we define

$$R(\tau_0^-) = R(0^-) := 0.$$

It is a similar standard observation that the process

$$\hat{X} = (\hat{X}(n) := (X(\tau_n), R_{\mathrm{S}}(\tau_n-)) : n \in \mathbb{N}_0)$$

is a homogeneous Markov chain on state space $\mathbb{N}_0 \times \{0\}$. Because of $R(\tau_n-) = 0\ \forall n \in \mathbb{N}_0$, we prefer to use for this Markov chain on state space $\mathbb{N}_0$ the description

$$\hat{X} = (\hat{X}(n) := X(\tau_n) : n \in \mathbb{N}_0).$$

A little reflection shows that the one-step transition matrix of $\hat{X}$ is a matrix which has the usual skip-free to the left property. The transition probabilities $\tilde{p}(i, n)$ are defined as

$$\tilde{p}(i, n) := P\left(X(\tau_1) = i + n - 1 | X(0) = i\right) \qquad \forall i \in \mathbb{N}_0, n \in \mathbb{N}. \tag{3.2.1}$$

Because

$$\tilde{p}(0, n) = \tilde{p}(1, n)$$

the one-step transition matrix of $\hat{X}$ is of the form (empty entries are zero)

$$\tilde{P} \;:=\; \begin{pmatrix} \tilde{p}(1,0) & \tilde{p}(1,1) & \tilde{p}(1,2) & \tilde{p}(1,3) & \dots \\ \tilde{p}(1,0) & \tilde{p}(1,1) & \tilde{p}(1,2) & \tilde{p}(1,3) & \dots \\ & \tilde{p}(2,0) & \tilde{p}(2,1) & \tilde{p}(2,2) & \dots \\ & & \tilde{p}(3,0) & \tilde{p}(3,1) & \dots \\ & & & & \end{pmatrix}, \tag{3.2.2}$$

which is an upper Hessenberg matrix. A similar one-step transition matrix arises in [Dsh97, p. 68] where the service requests are state dependent, but no speeds are incorporated.

So for $\tilde{P}$ the row index i indicates the number of customers in system when a service commences (and the service request is drawn according to G_n), and the (varying in row number) column index n indicates the number of customers who arrived during the ongoing service.

Note, that although we have used an intuitive notation for the non zero entries of $\tilde{P}$, the matrix is a fairly general upper Hessenberg matrix: The only restrictions are strict positivity of the $\tilde{p}(i,n)$ and row sum 1.

We define conditional probability for a system with i customer and residual service request r to admit n new customers until departure of the first customer

$$\begin{aligned} &\tilde{p}(i,n,r) := P\left(X(\tau_1) = i+n-1 | X(0) = i, R_{\mathrm{S}} = r\right), && \forall i \geq 1, n \in \mathbb{N}_0, r \in \mathbb{R}^+, \\ &\tilde{p}(i,n,0) := 0 && \forall i \geq 1, n \in \mathbb{N}_0. \end{aligned} \tag{3.2.3}$$

For $\tilde{p}(i,n)$ it holds

$$\tilde{p}(i,n) = \int_0^\infty \tilde{p}(i,n,r) dG_i(r). \tag{3.2.4}$$

We will not go further into the details of computing $\tilde{P}$, but will recall the classical result for state independent service speeds ($c(n) = 1$ for all $n \geq 1$) in the following subsection.

$M/G/1/\infty$ queueing systems The classical situation is as follows (See [Kle75, 176+]).

Proposition 3.2.1. *For the $M/G/1/\infty$ queuing system with service time distribution $G : [0,\infty) \to [0,1]$ and constant service speeds $c(n) \equiv 1$, the transition probabilities $\tilde{p}(i,n)$ are independent of i and the transition matrix $\tilde{P}$ has the form*

$$\tilde{P} := \begin{pmatrix} \tilde{p}(0) & \tilde{p}(1) & \tilde{p}(2) & \tilde{p}(3) & \dots \\ \tilde{p}(0) & \tilde{p}(1) & \tilde{p}(2) & \tilde{p}(3) & \dots \\ & \tilde{p}(0) & \tilde{p}(1) & \tilde{p}(2) & \dots \\ & & \tilde{p}(0) & \tilde{p}(1) & \dots \\ & & & & \end{pmatrix} \tag{3.2.5}$$

with

$$\tilde{p}(n) := \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t).$$

With $\mu^{-1} < \infty$ we denote the mean service time. Then under $\lambda\mu^{-1} < 1$ the continuous time process and the chain embedded at departure instants are ergodic.

We denote as usual the stationary distribution of $\hat{X}$ by $\hat{\xi}$. $\hat{\xi}$ is the unique stochastic solution of the equation

$$\hat{\xi}\tilde{P} = \hat{\xi} \tag{3.2.6}$$

and it is also equal to the steady-state distribution of the continuous time process X.

In this section, in Example 3.2.5, Example 3.2.6 and Proposition 3.2.8 we will analyze properties of systems with deterministic service times. We therefore recall well known results for standard $M/D/1/\infty$ queues where the service time is deterministic of length $\frac{1}{\mu}$, i.e., the distribution function is $G = \delta_{\frac{1}{\mu}}$ (Dirac measure). We assume $\rho := \lambda/\mu < 1$. Then the queue length process $\hat{X} = \left(\hat{X}(n) : n \in \mathbb{N}_0\right)$ at departure times is an ergodic Markov chain with one-step transition matrix (3.2.5) with

$$\tilde{p}(n) := \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} d\delta_{\frac{1}{\mu}}(t) = e^{-\frac{\lambda}{\mu}} \frac{(\frac{\lambda}{\mu})^n}{n!}. \tag{3.2.7}$$

We will utilize later on some special values of $\hat{\xi}$ (see [GH74, p. 241])

$$\hat{\xi}(0) = (1-\rho), \qquad \hat{\xi}(1) = (1-\rho)\left(e^\rho - 1\right), \qquad \hat{\xi}(2) = (1-\rho)\, e^\rho (e^\rho - \rho - 1). \tag{3.2.8}$$

3.2.2. $M/G/1/\infty$ system with inventory under lost sales

We analyze an $M/G/1/\infty$ queueing system with an attached inventory, reorder point r, random replenishment size depending on the inventory size at the replenishment time instant, general service times, and exponential replenishment time. See e.g to the special cases in [Sch04, Definition 1.3.16 and Definition 1.3.23], where service time is exponential and replenishment size depends on the the current policy: (r, Q) or (r, S). We summarize the system's parameters:

Poisson-λ input, service time is general with finite mean value $\frac{1}{\mu}$, $\rho := \lambda/\mu < 1$. Lead times are exponential-ν. All service, interarrival, and lead times constitute an independent family. When the inventory is depleted no service is provided and new arrivals are rejected (lost sales).

The Markovian state process of the integrated queueing-inventory system relies on the description of the $M/G/1/\infty$ queueing system, given at the beginning of Section 3.2.1.

For the system's description in continuous time we use the supplemented queue length process (X, R_{S}), where the process R_{S} on $\mathbb{R}_0^+$ denotes the residual service time of the ongoing service as the supplementary variable. We enlarge this process by adding the inventory size Y.

The joint queueing-inventory process with supplementary variable R_{S} will be denoted by $Z = (X, R_{\mathrm{S}}, Y)$, and lives on state space $\mathbb{N}_0 \times \mathbb{R}_0^+ \times K$. With $K = \{0, \ldots, \kappa\}$ where $\kappa \in \mathbb{N}$ is the maximal inventory level. We consider the system at departure instants, which leads to a one-step transition matrix similar to (3.2.5).

The dynamics of of the Markov chain $\hat{Z}$ embedded into Z at departure instants will be described in a way that resembles the $M/G/1$ type matrix analytical models.

From the structure of the embedding, we know, that $R_{\mathrm{S}}(\tau_n^-) = 0$ and whenever $X(\tau_n) = 0$ we see $R_{\mathrm{S}}(\tau_n) = 0$, resp. whenever $X(\tau_n) > 0$ we see $R_{\mathrm{S}}(\tau_n) \sim G$. We therefore can, without loss of information, delete the R-component of the process, to obtain a Markov chain embedded at departure times

$$\hat{Z} = (\hat{X}, \hat{Y}) = \big((\hat{X}(n), \hat{Y}(n)) : n \in \mathbb{N}_0\big), \text{ with } \hat{Z}(n) := (\hat{X}(n), \hat{Y}(n)) := (X(\tau_n), Y(\tau_n)).$$

The state space of $\hat{Z}$ is $E = \mathbb{N}_0 \times K$ where K is partitioned into $K = K_W \uplus K_B$ with $K_B = \{0\}$.

We proceed with nomenclature similar to Definition 3.1.3 with the obvious modifications, which stem from the observation, that for $i > 0$ the probabilities $P(Z(\tau_1) = (n+i-1, m) | Z(0) = (i, k))$ do not depend on i, because service is provided with an intensity which is independent of the queue length. We reuse several of the previous notations but there will be no danger of misinterpretation in this section. Recall, that $(\tau_n : n \in \mathbb{N}_0)$ is the sequence of departure instants

Definition 3.2.2. We introduce matrices $A^{(n)} \in \mathbb{R}^{K \times K}$ by

$$A_{km}^{(n)} := P(Z(\tau_1) = (n+i-1, m) | Z(0) = (i, k)), \qquad 1 \le i, \tag{3.2.9}$$

for $k, m \in K$. Then the one-step transition matrix $\mathbf{P}$ defined according to Definition 3.1.3 has the form

$$\mathbf{P} = \begin{pmatrix} B^{(0)} & B^{(1)} & B^{(2)} & B^{(3)} & \dots \\ A^{(0)} & A^{(1)} & A^{(2)} & A^{(3)} & \dots \\ 0 & A^{(0)} & A^{(1)} & A^{(2)} & \dots \\ 0 & 0 & A^{(0)} & A^{(1)} & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}. \tag{3.2.10}$$

We will clarify the structure of the solution of the equation $\hat{\pi}\mathbf{P} = \hat{\pi}$. So, $\hat{\pi}$ is the steady-state distribution of the embedded Markov chain $\hat{Z}$. It will become clear that $\hat{Z}$ is in general not irreducible on E.

As in Definition 3.1.2 we will call the marginal steady-state distributions

$$\hat{\xi} \quad := \quad (\hat{\xi}(n) : n \in \mathbb{N}_0), \qquad \hat{\xi}(n) := \sum_{k \in K} \hat{\pi}(n,k). \tag{3.2.11}$$

and

$$\hat{\theta} \quad := \quad (\hat{\theta}(k) : k \in K), \qquad \hat{\theta}(k) := \sum_{n \in \mathbb{N}_0} \hat{\pi}(n,k).$$

It will be convenient to group $\hat{\pi}$ as

$$\hat{\pi} = (\hat{\pi}^{(0)}, \hat{\pi}^{(1)}, \hat{\pi}^{(2)}, \dots) \tag{3.2.12}$$

with

$$\hat{\pi}^{(n)} = (\hat{\pi}(n,0), \hat{\pi}(n,1), \dots, \hat{\pi}(n,S)), \qquad n \in \mathbb{N}_0. \tag{3.2.13}$$

An immediate consequence is that the steady-state equation can be written as

$$\hat{\pi}^{(0)} B^{(n)} + \sum_{i=1}^{n+1} \hat{\pi}^{(i)} A^{(n-i+1)} = \hat{\pi}^{(n)}, \qquad n \in \mathbb{N}_0. \tag{3.2.14}$$

We determine $A^{(n)}$, $B^{(n)}$ explicitly, distinguishing cases by the initial states $\hat{Z}(0)$ according to (3.2.9).

Definition 3.2.3. Consider an $M/G/1/\infty$ system with reorder point r, and random replenishment size and lost sales. The environment states represent inventory size. We have an environment state set $K = \{0, \dots, \kappa\}$ where κ is the maximal size of the inventory. The replenishment time is exponential with rate ν. The replenishment order is immediately triggered, as soon as the inventory reaches or falls below r. The replenishment order is placed as long as the inventory level stays below or equal to r. When the inventory is empty no service is provided and new customers are lost. The arrival stream is Poisson-λ and service time distribution G is general with mean value $\frac{1}{\mu}$ and $\lambda < \mu$.

We assume that the probability for the inventory to change from state k to state m after replenishment is $s(k,m)$, where $s := (s(k,m) : k, m \in K)$ is a stochastic matrix.

We assume that after replenishment we have more items in the stock, so

$$s(k,m) = 0 \text{ for } m \le k$$

and that the inventory size after replenishment is greater than the reorder point r, i.e.

$$s(k,m) = 0 \text{ for } m \le r.$$

The transition probability of the embedded Markov chains of the system are determined as follows:

- $\hat{Z}(0) = (i, 0),\ i > 0$: The server waits for replenishment of inventory. The queue length stays at i until the ordered replenishment arrives. Then the inventory is restocked to $m > 0$ with probability $s(0, m)$ and the server resumes work, stochastically identical to a standard $M/G/1/\infty$-system until the service expires. When the served customer leaves the system, the inventory contains $m - 1$ item.

$$\begin{aligned} A^{(n)}_{(0,m-1)} &= P(Z(\tau_1) = (n + i - 1, m - 1)|Z(0) = (i, 0)) \\ &= s(0, m) \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) = s(0, m)\tilde{p}(n). \end{aligned}$$

- $\hat{Z}(0) = (i, k),\ i > 0,\ 1 \leq k \leq r$: A lead time is ongoing and the server is active serving the first customer in the queue. In this case there are two possible target state groups for the inventory when the customer currently in service leaves the system.
  - *Target state* $k - 1$: The ongoing service expires before the lead time does. The resulting inventory state after service is finished is $k - 1$.

$$\begin{aligned} A^{(n)}_{(k,k-1)} &= P(Z(\tau_1) = (n + i - 1, k - 1)|Z(0) = (i, k)) \\ &= \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} dG(t) \end{aligned} \tag{3.2.15}$$

  - *Target state* $m - 1 \geq k$: The ongoing lead expires before the service time does, and the inventory is filled up to m during the ongoing service. The resulting inventory state when service expired is $m - 1$. (If $m - 1 \leq r$, an additional order is placed, but this does not change the state.)

$$\begin{aligned} A^{(n)}_{(k,m-1)} &= P(Z(\tau_1) = (n + i - 1, m - 1)|Z(0) = (i, k)) \\ &= s(k, m) \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} (1 - e^{-\nu t}) dG(t) \\ &= s(k, m) \left(\tilde{p}(n) - \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} dG(t) \right) \end{aligned}$$

- $\hat{Z}(0) = (i, k),\ i > 0,\ r + 1 \leq k$: There are k items on stock, no order is placed and the service is provided just as in a standard $M/G/1/\infty$ system. The resulting inventory state when service expired is $k - 1$.

$$\begin{aligned} A^{(n)}_{(k,k-1)} &= P(Z(\tau_1) = (n + i - 1, k - 1)|Z(0) = (i, k)) \\ &= \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) = \tilde{p}(n) \end{aligned}$$

- $\hat{Z}(0) = (0, 0)$: The queue is empty, an order is placed. No customers are admitted until replenishment of inventory. When the ongoing lead time expires, inventory is restocked to $m \geq r$ with probability $s(0, m)$. Thereafter new customers are admitted, and service starts immediately after the first arrival. When this customer is served, the stock size is $m-1 \geq r-1$.

$$\begin{aligned} B^{(n)}_{(0,m-1)} &= P(Z(\tau_1) = (n, m-1)|Z(0) = (0,0)) \\ &= s(0,m) \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) = s(0,m)\tilde{p}(n) \end{aligned}$$

- $\hat{Z}(0) = (0,k)$, $1 \leq k \leq r$: The queue is empty, there are k items on stock, and an order is placed. In this case there are two possible target states for the inventory when the first customer who arrives is served and leaves the system.
  - *Target state* $k-1$: The ongoing inter-arrival time expires before the lead time does. The arriving customer's service starts immediately and is finished before the replenishment arrives. The resulting inventory state after service is finished is $k-1$.

$$\begin{aligned} B^{(n)}_{(k,k-1)} &= P(Z(\tau_1) = (n, k-1)|Z(0) = (0,k)) \\ &= \frac{\lambda}{\nu+\lambda} \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} dG(t) \end{aligned} \tag{3.2.16}$$

  - *Target state* $m-1 \geq r-1$:
    1. The ongoing lead expires before the inter-arrival time does, and the inventory is filled up to m with probability $s(k,m)$ during the ongoing inter-arrival time. Then, until the first departure, the system acts like a standard $M/G/1/\infty$ queue. When the first departure happens, inventory size decreases to $m-1$.
    2. The ongoing inter-arrival time expires before the lead time does. The arriving customer's service starts immediately and the replenishment arrives before the service is finished and by the replenishment the stock size increases to m. The resulting inventory state after service is finished is $m-1$.

$$\begin{aligned} B^{(n)}_{(k,m-1)} =& P(Z(\tau_1) = (n,1)|Z(0) = (0,k)) \\ =& s(k,m) \frac{\nu}{\nu+\lambda} \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) \\ &+ s(k,m) \frac{\lambda}{\nu+\lambda} \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} (1-e^{-\nu t}) dG(t) \\ =& s(k,m) \left(\tilde{p}(n) - \frac{\lambda}{\nu+\lambda} \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} dG(t) \right) \end{aligned}$$

- $\hat{Z}(0) = (0,k)$, $r+1 \leq k$: The queue is empty, there are k items on stock, and an inter-arrival time is ongoing. Until the first departure the system develops like a standard $M/G/1/\infty$ queue. After that departure the inventory size is $k-1$.

$$B^{(n)}_{(k,k-1)} = P(Z(\tau_1) = (n, k-1)|Z(0) = (0,k)) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(t) = \tilde{p}(n)$$

The remaining transition probabilities are zero.

Lemma 3.2.4. *The marginal steady-state distribution of the number of customers in the $M/G/1/\infty$ system with inventory and random replenishment size as defined in Definition 3.2.3 is equal to the distribution of the number of the customers in a $M/G/1/\infty$ system with the same arrival rate and service time distribution as defined in Proposition 3.2.1. It is unique and exists if and only if $\lambda < \mu$.*

Proof. From the Definition 3.2.3 it follows for all $n \in \mathbb{N}_0$:

$$\begin{aligned}
&\sum_{m\in K} B^{(n)}_{(0,m)} = \sum_{m\in K} s(0,m)\tilde{p}(n) = \tilde{p}(n)\\
\text{for } 1 \le k \le r: \quad &\sum_{m\in K} B^{(n)}_{(k,m)} = B^{(n)}_{(k,k-1)} + \sum_{m\ge k} B^{(n)}_{(k,m)}\\
&= \frac{\lambda}{\nu+\lambda}\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}dG(t)\\
&\quad + \underbrace{\sum_{m>k} s(k,m)}_{=1}\left(\tilde{p}(n) - \frac{\lambda}{\nu+\lambda}\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}dG(t)\right)\\
&= \tilde{p}(n), \qquad 1 \le k \le r,\\
\text{for } r+1 \le k: \quad &\sum_{m\in K} B^{(n)}_{(k,m)} = B^{(n)}_{(k,k-1)} = \tilde{p}(n), \qquad r+1 \le k.
\end{aligned}$$

Summarizing the equations above, the row sums of the matrix $B^{(n)}$ are equal to $\tilde{p}(n)$:

$$B^{(n)}\mathbf{e} = \tilde{p}(n)\mathbf{e}. \tag{3.2.17}$$

Similar we show that the row sums of the matrix $A^{(n)}$ are $\tilde{p}(n)$:

$$\begin{aligned}
&\sum_{m\in K} A^{(n)}_{(0,m)} = \sum_{m\in K} s(0,m)\tilde{p}(n) = \tilde{p}(n)\\
\text{for } 1 \le k \le r: \quad &\sum_{m\in K} A^{(n)}_{(k,m)} = A^{(n)}_{(k,k-1)} + \sum_{m\ge k} A^{(n)}_{(k,m)}\\
&= \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}dG(t)\\
&\quad + \underbrace{\sum_{m\ge k} s(k,m)}_{=1}\left(\tilde{p}(n) - \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}dG(t)\right)\\
&= \tilde{p}(n), \qquad 1 \le k \le r,\\
\text{for } r+1 \le k: \quad &\sum_{m\in K} A^{(n)}_{(k,m)} = A^{(n)}_{(k,k-1)} = \tilde{p}(n), \qquad r+1 \le k,
\end{aligned}$$

$$\Longrightarrow A^{(n)}\mathbf{e} = \tilde{p}(n)\mathbf{e}. \tag{3.2.18}$$

Multiplying the steady-state equations (3.2.14) for $\hat{Z}$ from the right with $\mathbf{e}$ and using the row sum properties (3.2.17) and (3.2.18) lead to

$$\hat{\pi}^{(0)}B^{(n)}\mathbf{e}+\sum_{i=1}^{n+1}\hat{\pi}^{(i)}A^{(i,n-i+1)}\mathbf{e}=\hat{\pi}^{(n)}\mathbf{e}\Longrightarrow\hat{\pi}^{(0)}\mathbf{e}\tilde{p}(0)+\sum_{i=1}^{n+1}\hat{\pi}^{(i)}\mathbf{e}\tilde{p}(n-i+1)=\hat{\pi}^{(n)}\mathbf{e}.$$

Using definition of $\hat{\xi}$ in (3.2.11) for $\hat{\pi}^{(i)}\mathbf{e}=\sum_{k\in K}\hat{\pi}(i,k)=\hat{\xi}(i)$ the last equation can be written as

$$\hat{\xi}(0)\tilde{p}(0)+\sum_{i=1}^{n+1}\hat{\xi}(i)\tilde{p}(n+1-i)=\hat{\xi}(n)$$

which is the steady-state equation (3.2.6) of a $M/G/1/\infty$ queue without environment in Proposition 3.2.1. Because transition probabilities $\tilde{p}(n)$ are positive for all $\mathbb{N}_0$, this equation has a unique up to a constant factor solution. This solution can be normalized if and only if $\lambda<\mu$. $\square$

Example 3.2.5. $M/D/1/\infty$ system with (r,S) policy and lost sales. It is a special case of the system from Definition 3.2.3 with deterministic service time distribution, $\kappa=S$, and $s(k,S)=1$ for $0\le k\le r$. The positive transition probabilities are:

$$A^{(n)}_{(0,S-1)}=\tilde{p}(n),$$

$$A^{(n)}_{(k,k-1)}=\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}d\delta_{\frac{1}{\mu}}(t)=e^{-\frac{\lambda+\nu}{\mu}}\frac{(\frac{\lambda}{\mu})^n}{n!}=e^{-\frac{\nu}{\mu}}\tilde{p}(n),\qquad 1\le k\le r,$$

$$\begin{aligned}A^{(n)}_{(k,S-1)}&=\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}(1-e^{-\nu t})d\delta_{\frac{1}{\mu}}(t)=(1-e^{-\frac{\nu}{\mu}})e^{-\frac{\lambda}{\mu}}\frac{(\frac{\lambda}{\mu})^n}{n!}\\&=(1-e^{-\frac{\nu}{\mu}})\tilde{p}(n),\qquad 1\le k\le r,\end{aligned}$$

$$A^{(n)}_{(k,k-1)}=\tilde{p}(n),\qquad r+1\le k,$$

$$B^{(n)}_{(0,S-1)}=\tilde{p}(n),$$

$$B^{(n)}_{(k,k-1)}=\frac{\lambda}{\nu+\lambda}\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}d\delta_{\frac{1}{\mu}}(t)=\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}}\tilde{p}(n),\qquad 1\le k\le r,$$

$$\begin{aligned}B^{(n)}_{(k,S-1)}&=\tilde{p}(n)-\frac{\lambda}{\nu+\lambda}\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}d\delta_{\frac{1}{\mu}}(t)\\&=\left(1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}}\right)\tilde{p}(n),\qquad 1\le k\le r,\end{aligned}$$

$$B^{(n)}_{(k,k-1)}=\tilde{p}(n),\qquad r+1\le k.$$

Summarizing the results we have:

$$A^{(n)} = \tilde{p}(n)\left(\begin{array}{c|ccccccccc}
 & 0 & 1 & \dots & r-1 & r & r+1 & \dots & S-1 & S \\ \hline
0 & 0 & 0 & & 0 & 0 & & & 1 & 0 \\
1 & e^{-\frac{\nu}{\mu}} & 0 & & & & & & 1-e^{-\frac{\nu}{\mu}} & 0 \\
\vdots & & \ddots & \ddots & & & & & \vdots & \\
r-1 & 0 & & \ddots & \ddots & & & & 1-e^{-\frac{\nu}{\mu}} & \\
r & 0 & & & e^{-\frac{\nu}{\mu}} & 0 & & & 1-e^{-\frac{\nu}{\mu}} & \\
r+1 & 0 & & & & 1 & 0 & & 0 & \\
\vdots & & & & & & \ddots & \ddots & & \\
S-1 & 0 & & & & & & \ddots & 0 & 0 \\
S & 0 & & & & & & & 1 & 0
\end{array}\right)$$

and

$$B^{(n)} = \tilde{p}(n)\left(\begin{array}{c|ccccccccc}
 & 0 & 1 & \dots & r-1 & r & r+1 & \dots & S-1 & S \\ \hline
0 & 0 & 0 & & 0 & 0 & & & 1 & 0 \\
1 & \frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & 0 & & & & & & 1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & 0 \\
\vdots & & \ddots & \ddots & & & & & \vdots & \\
r-1 & 0 & & \ddots & \ddots & & & & 1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & \\
r & 0 & & & \frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & 0 & & & 1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & \\
r+1 & 0 & & & & 1 & 0 & & 0 & \\
\vdots & & & & & & \ddots & \ddots & & \\
S-1 & 0 & & & & & & \ddots & 0 & 0 \\
S & 0 & & & & & & & 1 & 0
\end{array}\right).$$

In the special case $r+1=S$, it holds

$$A^{(n)} = \tilde{p}(n)\left(\begin{array}{c|ccccc}
 & 0 & 1 & r-1 & r & S \\ \hline
0 & 0 & 0 & 0 & 1 & 0 \\
1 & e^{-\frac{\nu}{\mu}} & & & 1-e^{-\frac{\nu}{\mu}} & 0 \\
 & & \ddots & & \vdots & \\
r-1 & & & \ddots & 1-e^{-\frac{\nu}{\mu}} & \\
r & & & e^{-\frac{\nu}{\mu}} & 1-e^{-\frac{\nu}{\mu}} & 0 \\
S & & & & 1 & 0
\end{array}\right),$$

$$B^{(n)} = \tilde{p}(n)\left(\begin{array}{c|cccccc}
 & 0 & 1 & \dots & r-1 & r & S \\ \hline
0 & 0 & 0 & & 0 & 1 & 0 \\
1 & \frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & 0 & & & 1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & 0 \\
\vdots & & \ddots & \ddots & & \vdots & \\
r-1 & 0 & & \ddots & \ddots & 1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & \\
r & 0 & & & \frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & 1-\frac{\lambda}{\nu+\lambda}e^{-\frac{\nu}{\mu}} & \\
S & 0 & & & & 1 & 0
\end{array}\right).$$

Example 3.2.6. $M/D/1/\infty$ system with (r, Q) policy and lost sales. It is a special case of a system in Definition 3.2.3 with deterministic service time distribution, $\kappa = r + Q$, and $s(k, k + Q) = 1$ for $0 \le k \le r$. The positive transition probabilities are:

$$A^{(n)}_{(0,Q-1)} = \tilde{p}(n),$$

$$A^{(n)}_{(k,k-1)} = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} d\delta_{\frac{1}{\mu}}(t) = e^{-\frac{\lambda+\nu}{\mu}} \frac{(\frac{\lambda}{\mu})^n}{n!} = e^{-\frac{\nu}{\mu}} \tilde{p}(n), \qquad 1 \le k \le r,$$

$$\begin{aligned} A^{(n)}_{(k,k+Q-1)} &= \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} (1 - e^{-\nu t}) d\delta_{\frac{1}{\mu}}(t) = (1 - e^{-\frac{\nu}{\mu}}) e^{-\frac{\lambda}{\mu}} \frac{(\frac{\lambda}{\mu})^n}{n!} \\ &= (1 - e^{-\frac{\nu}{\mu}}) \tilde{p}(n), \qquad 1 \le k \le r, \end{aligned}$$

$$A^{(n)}_{(k,k-1)} = \tilde{p}(n), \qquad r + 1 \le k,$$

$$B^{(n)}_{(0,Q-1)} = \tilde{p}(n),$$

$$B^{(n)}_{(k,k-1)} = \frac{\lambda}{\nu + \lambda} \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} d\delta_{\frac{1}{\mu}}(t) = \frac{\lambda}{\nu + \lambda} e^{-\frac{\nu}{\mu}} \tilde{p}(n), \qquad 1 \le k \le r,$$

$$\begin{aligned} B^{(n)}_{(k,k+Q-1)} &= \tilde{p}(n) - \frac{\lambda}{\nu + \lambda} \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\nu t} d\delta_{\frac{1}{\mu}}(t) \\ &= \left(1 - \frac{\lambda}{\nu + \lambda} e^{-\frac{\nu}{\mu}}\right) \tilde{p}(n), \qquad 1 \le k \le r, \end{aligned}$$

$$B^{(n)}_{(k,k-1)} = \tilde{p}(n), \qquad r + 1 \le k.$$

Summarizing the results we have:

$$A^{(n)} =$$

$$\tilde{p}(n) \left(\begin{array}{c|cccc} & 0,\dots,r-1 & r,\dots,Q-1 & Q,\dots,r+Q-1 & r+Q \\ \hline 0 & 0 & 0,\dots,0,1 & & 0 \\ \begin{matrix} 1 \\ \vdots \\ r \end{matrix} & \begin{pmatrix} e^{-\frac{\nu}{\mu}} & & \\ & \ddots & \\ & & e^{-\frac{\nu}{\mu}} \end{pmatrix} & & I - \begin{pmatrix} e^{-\frac{\nu}{\mu}} & & \\ & \ddots & \\ & & e^{-\frac{\nu}{\mu}} \end{pmatrix} & 0 \\ \begin{matrix} r+1 \\ \vdots \\ Q \end{matrix} & 0 & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} & & 0 \\ \begin{matrix} Q+1 \\ \vdots \\ r+Q \end{matrix} & 0 & & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} & 0 \end{array} \right)$$

and

$$B^{(n)} =$$

$$\tilde{p}(n) \left(\begin{array}{c|cccc} & 0\dots r-1 & r\dots Q-1 & Q\dots r+Q-1 & r+Q \\ \hline 0 & 0 & 0\dots 0,1 & & 0 \\ \begin{matrix} 1 \\ \vdots \\ r \end{matrix} & \frac{\lambda}{\nu+\lambda}\begin{pmatrix} e^{-\frac{\nu}{\mu}} & & \\ & \ddots & \\ & & e^{-\frac{\nu}{\mu}} \end{pmatrix} & & I - \frac{\lambda}{\nu+\lambda}\begin{pmatrix} e^{-\frac{\nu}{\mu}} & & \\ & \ddots & \\ & & e^{-\frac{\nu}{\mu}} \end{pmatrix} & 0 \\ \begin{matrix} r+1 \\ \vdots \\ Q \end{matrix} & 0 & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} & & 0 \\ \begin{matrix} Q+1 \\ \vdots \\ r+Q \end{matrix} & 0 & & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} & 0 \end{array} \right).$$

Remark 3.2.7. The systems in Example 3.2.5 and Example 3.2.6 describe special cases of $M/G/1/\infty$ systems in [Vin08, Section 5.3] when service time distribution is deterministic. Example 3.2.6 is a special case for $M/G/1/\infty$ system with deterministic service time and parameter $\gamma = 1$ in [KML13, Section 6].

Proposition 3.2.8. *There exist parameters λ, μ, and ν such that any $M/D/1/\infty$ inventory with these parameters, reorder point $r \geq 1$, and random replenishment size greater than 1 – i.e. $s(k,0) = s(k,1) = 0$ for all k – does not have a product-form stationary distribution for the embedded Markov chains at departure times.*

Proof. We use the results from (3.2.16) and (3.2.15)

$$B^{(n)}_{(1,0)} = \frac{\lambda}{\nu+\lambda}\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}d\delta_{\frac{1}{\mu}}(t) = \frac{\lambda}{\nu+\lambda}e^{-\frac{\lambda+\nu}{\mu}}\left(\frac{\lambda}{\mu}\right)^n\frac{1}{n!},$$

$$A^{(n)}_{(1,0)} = \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}e^{-\nu t}d\delta_{\frac{1}{\mu}}(t) = e^{-\frac{\lambda+\nu}{\mu}}\left(\frac{\lambda}{\mu}\right)^n\frac{1}{n!},$$

and $B^{(n)}_{(k,0)} = A^{(n)}_{(k,0)} = 0$ for $k \neq 1$ to construct a system with a non-product-form steady-state distribution.

According to Lemma 3.2.4 the marginal distribution of the queue size is $\hat{\xi}$. Inserting this product form $\hat{\pi}(n,k) = \hat{\xi}(n)\hat{\theta}(k)$ into the equation for the level $n = 0$ and phase $k = 0$, the steady-state equation (3.2.14) is transformed into

$$\begin{aligned} \sum_{k\in K}\hat{\pi}(0,k)B^{(0)}_{(k,0)} + \sum_{k\in K}\hat{\pi}(1,k)A^{(0)}_{(k,0)} &= \hat{\pi}(0,0) \\ \Longleftrightarrow \hat{\pi}(0,1)B^{(0)}_{(1,0)} + \hat{\pi}(1,1)A^{(0)}_{(1,0)} &= \hat{\pi}(0,0) \\ \Longleftrightarrow \frac{\lambda}{\nu+\lambda}e^{-\frac{\lambda+\nu}{\mu}}\hat{\xi}(0)\hat{\theta}(1) + e^{-\frac{\lambda+\nu}{\mu}}\hat{\xi}(1)\hat{\theta}(1) &= \hat{\xi}(0)\hat{\theta}(0) \\ \Longleftrightarrow e^{-\frac{\lambda+\nu}{\mu}}\left(\frac{\lambda}{\nu+\lambda} + \frac{\hat{\xi}(1)}{\hat{\xi}(0)}\right)\hat{\theta}(1) &= \hat{\theta}(0). \end{aligned}$$

Substituting the values for $\hat{\xi}(0)$ and $\hat{\xi}(1)$ from (3.2.8) yields

$$\Longleftrightarrow e^{-\frac{\lambda+\nu}{\mu}}\left(\frac{\lambda}{\nu+\lambda} + e^{\rho} - 1\right)\hat{\theta}(1) = \hat{\theta}(0). \tag{3.2.19}$$

The equation for level $n = 1$ and phase $k = 0$ under this product form assumption is transformed into

$$\begin{aligned}
&\sum_{k\in K}\hat{\pi}(0,k)B^{(1)}_{(k,0)} + \sum_{k\in K}\hat{\pi}(1,k)A^{(1)}_{(k,0)} + \sum_{k\in K}\hat{\pi}(2,k)A^{(0)}_{(k,0)} = \hat{\pi}(1,0)\\
&\iff \hat{\pi}(0,1)B^{(1)}_{(1,0)} + \hat{\pi}(1,1)A^{(1)}_{(1,0)} + \hat{\pi}(2,1)A^{(0)}_{(1,0)} = \hat{\pi}(1,0)\\
&\iff \hat{\xi}(0)\hat{\theta}(1)B^{(1)}_{(1,0)} + \hat{\xi}(1)\hat{\theta}(1)A^{(1)}_{(1,0)} + \hat{\xi}(2)\hat{\theta}(1)A^{(0)}_{(1,0)} = \hat{\xi}(1)\hat{\theta}(0)\\
&\iff \left(\frac{\lambda}{\nu+\lambda}e^{-\frac{\lambda+\nu}{\mu}}\frac{\lambda}{\mu}\hat{\xi}(0) + e^{-\frac{\lambda+\nu}{\mu}}\frac{\lambda}{\mu}\hat{\xi}(1) + e^{-\frac{\lambda+\nu}{\mu}}\hat{\xi}(2)\right)\hat{\theta}(1) = \hat{\xi}(1)\hat{\theta}(0)\\
&\iff e^{-\frac{\lambda+\nu}{\mu}}\left(\frac{\lambda}{\nu+\lambda}\frac{\lambda}{\mu}\frac{\hat{\xi}(0)}{\hat{\xi}(1)} + \frac{\lambda}{\mu} + \frac{\hat{\xi}(2)}{\hat{\xi}(1)}\right)\hat{\theta}(1) = \hat{\theta}(0).
\end{aligned}$$

Substituting the values for $\hat{\xi}(0)$, $\hat{\xi}(1)$, and $\hat{\xi}(2)$ from (3.2.8) yields

$$\iff e^{-\frac{\lambda+\nu}{\mu}}\left(\frac{\lambda}{\nu+\lambda}\frac{\lambda}{\mu}\frac{1}{e^{\rho}-1} + \frac{\lambda}{\mu} + \frac{e^{\rho}(e^{\rho}-\rho-1)}{(e^{\rho}-1)}\right)\hat{\theta}(1) = \hat{\theta}(0). \tag{3.2.20}$$

One can see that the expressions (3.2.19) and (3.2.20) are in general not compatible. For example, with the parameters $\lambda = 1$, $\mu = 2$ and $\nu = 3$ the $\hat{\theta}(0)$ from (3.2.19) is approximately $0.122\cdot\hat{\theta}(1)$ and the $\hat{\theta}(0)$ from the (3.2.20) is approximately $0.145\cdot\hat{\theta}(1)$.

This result differs from product properties of $M/G/1/\infty$ inventory systems in [Vin08, Theorem 5.3.1 and Theorem 5.3.2] when $r \geq 1$.

□

3.2.3. $M/G/1/\infty$ queueing systems with state dependent service intensities and product-form steady state

In the previous section we have shown by a counterexample, that in general the steady-state distribution of an $M/G/1/\infty$ system with (r, S) policy and lost sales does not have a product form. Nevertheless, there are cases where loss systems with non-exponential service times in a random environment have product form steady states. These systems belong to a class of generalized $M/G/1/\infty$ loss systems, which will be discussed in this subsection.

Definition 3.2.9. We consider an $M/G/1/\infty$ queueing system in continuous time with state dependent service intensities (speeds) as described at the beginning of Section 3.2.1 (page 109) and use the notation introduced there.

The supplemented queue length process (X, R_{S}) (queue length, residual service request) is not Markov because we additionally assume that this queueing system is coupled with a **finite** environment $K = K_W \uplus K_B$ with $K_W \neq \emptyset$, driven again by a generator V and a stochastic jump matrix R, as described at the beginning of Section 2.1.1. The state of the environment process will be denoted by Y again.

We prescribe that the interaction of (X, R_{S}) with the environment process Y is via the following principles and restrictions:

(1) If the environment process is in a *non-blocking* state k, i.e. $k \in K_W$, the queueing system develops in the same way as an $M/G/1/\infty$ queuing system in isolation, governed by $\tilde{P}$ from (3.2.2), without any change of the environment until the next departure happens. Holding the environment invariant during this period is guaranteed by $v(k,m) = 0$ for all $k \in K_W, m \in K$.

(2) If at time t a customer departs from the system, the environment state changes according to the stochastic jump matrix R, independent of the history of the system given $Y(t)$.

(3) Whenever the environment process is in a *blocking* state $k \in K_B$, it may change its state with rates governed by the matrix V, independent of the queue length and the residual service request.

From these assumptions it is immediate, that $Z = (X, R_{\mathrm{S}}, Y)$ is a continuous time strong Markov process. We introduce sequences of stopping times for the process $Z = (X, R_{\mathrm{S}}, Y)$ as before: With $\tau_0 = \sigma_0 = \zeta_0 = 0$ we will denote by
$\tau = (\tau_0, \tau_1, \dots)$ the sequence of departure times of customers,
$\sigma = (\sigma_0, \sigma_1, \dots)$ the sequence of arrival times of customers admitted to the system,
$\zeta = (\zeta_0, \zeta_1, \dots)$ the sequence of jump times of the continuous time process Z.

By standard arguments it is seen that the sequence

$$(X(\tau_n), R_{\mathrm{S}}(\tau_n^-), Y(\tau_n)) : n \in \mathbb{N}_0)$$

is a homogeneous Markov chain on state space $\mathbb{N}_0 \times \{0\} \times K$. Consider henceforth the homogeneous Markov chain

$$\hat{Z} = (\hat{X}, \hat{R}_{\mathrm{S}}, \hat{Y}) = \big((\hat{X}(n), \hat{R}_{\mathrm{S}}(n), \hat{Y}(n)) : n \in \mathbb{N}_0\big)$$

on state space $\mathbb{N}_0 \times \{0\} \times K$ with

$$\hat{Z}(n) := (\hat{X}(n), \hat{R}_{\mathrm{S}}(n), \hat{Y}(n)) := (X(\tau_n), R_{\mathrm{S}}(\tau_n^-), Y(\tau_n)), \qquad \forall n \in \mathbb{N}_0.$$

Because for all $n \in \mathbb{N}_0$ it holds $R_{\mathrm{S}}(\tau_n^-) = 0$, we can later omit the $\hat{R}_{\mathrm{S}}$-component of $\hat{Z}$. Nevertheless, for the intermediate system analysis and for many proofs of this section, we will trace the value $\hat{R}_{\mathrm{S}}$ for vividness.

The following formulae follow directly from the description.

(1) $\Longrightarrow$ for $k \in K_W,\ m \in K$

$$P\Big(\big(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1^-)\big) = (i+n-1, 0, m) \Big| \big(X(0), R_{\mathrm{S}}(0^-), Y(0)\big) = (i, 0, k)\Big)$$
$$= \delta_{km} \tilde{p}(i,n).$$

(2) $\Longrightarrow$ for $k \in K_W$, $m \in K$

$$P\Big(\big(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\big) = (i+n-1, 0, m) \Big| \big(X(0), R_{\mathrm{S}}(0^-), Y(0)\big) = (i, 0, k)\Big)$$
$$= \sum_{h \in K} P\Big(\big(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1^-)\big) = (i+n-1, 0, h) \Big| \big(X(0), R_{\mathrm{S}}(0^-), Y(0)\big) = (i, 0, k)\Big)$$
$$\cdot R(h,m)$$
$$= \tilde{p}(i,n) \cdot R(k,m)$$

(3) $\Longrightarrow$ for $i \in \mathbb{N}_0$, $k \in K_B$, $m \in K$, and $r, s \in \mathbb{R}_0^+$

$$P\Big((X(\zeta_1), R_{\mathrm{S}}(\zeta_1), Y(\zeta_1)) = (j,s,m)\Big|(X, R_{\mathrm{S}}, Y)(0) = (i,r,k)\Big) = \delta_{ij}\delta_{rs}\frac{v(k,m)}{-v(k,k)}.$$

Note, that in the last expression $k \in K_B$ implies that the queueing system is frozen, and therefore in the denominator of the right side no summands originating from arrival or service process occur.

Although we have imposed constraints on the behaviour of the environment the model still is a very versatile one. The class of models from Definition 3.2.9 encompasses (e.g.) many vacation models. These are models describing a server working on primary and secondary customers, a situation which arises in many computer, communication, and production systems and networks. If one is mainly interested in the service process of primary customers, then working on secondary customers means from the viewpoint of the primary customers, that the server is not available or is interrupted. For more details see e.g. the survey of Doshi [Dos90]. In the classification given there [Dos90, pp. 221, 222] the above model is a single server queue with *general nonexhaustive service, with non-preemptive vacations,* and *general vacation rule.* Our system fits into these classification because whenever a service expires the server decides to take a vacation for serving secondary customers (a state K_B is selected) or to continue to serve a customer if there is any (a state in K_W is selected).

The proposed product form property of $\hat{Z}$ originates from the specific structure of the one-step transition matrix $\mathbf{P}$ of $(\hat{X}, \hat{Y})$. With some stochastic matrix $H \in \mathbb{R}^{K\times K}$, which we present in all details below,

$$\mathbf{P} = \begin{pmatrix} \tilde{p}(1,0)H & \tilde{p}(1,1)H & \tilde{p}(1,2)H & \tilde{p}(1,3)H & \dots \\ \tilde{p}(1,0)H & \tilde{p}(1,1)H & \tilde{p}(1,2)H & \tilde{p}(1,3)H & \dots \\ 0 & \tilde{p}(2,2)H & \tilde{p}(2,1)H & \tilde{p}(2,2)H & \dots \\ 0 & 0 & \tilde{p}(3,0)H & \tilde{p}(3,1)H & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}. \tag{3.2.21}$$

We will use an evaluation procedure similar to that used for the $M/M/1/\infty$ in a random environment, by decomposing the matrices $B^{(n)} = WU^{(n,0)}R$ and $A^{(i,n)} = U^{(i,n)}R$.

Assumption 3.2.10. *To avoid discussion of ergodic theory for processes with states in continuous space we also assume for Section 3.2.3:*

(i) The probabilities $P(X(t) = n, R_S(t) \le r, Y(t) = k)$ *converge to unique limiting and steady-state probabilities* $\lim_{t\to\infty} P(X(t) = n, R_S(t) \le r, Y(t) = k)$ *independent from the starting probability.*

(ii) It holds $\lim_{t\to\infty} P(X(t) = n, Y(t) = k) > 0$.

(iii) The steady-state probability of the embedded Markov chain $(\hat{X}(n), \hat{Y}(n))$ *exists and it is unique.*

Further to simplify the proof we assume that

(iv) The steady-state distribution $\hat{\xi}$ *of the pure queuing process, with the same service time and arrival time parameters as the system from Definition 3.2.9 exists. It holds* $\hat{\xi}\tilde{P} = \hat{\xi}$ *with one-step transition matrix* (3.2.2).

Remark 3.2.11. In the case of phase-type distributed service time with finite number L of phases, Assumption 3.2.10 (i)–(iii) follow from ergodicity assumption.

The next lemma guarantees that the expressions $\frac{1}{-v(k,k)+1_{[k\in K_W]}}$ and $(I_W - V)^{-1}$ in Lemma 3.2.14 and Lemma 3.2.18 are always well defined.

Lemma 3.2.12. *For the system defined in Definition 3.2.9 it holds:*

(i)

$$|v(k,k)| > 0, \qquad \forall k \in K_B. \tag{3.2.22}$$

Therefore the expression

$$\frac{1}{-v(k,k) + 1_{[k\in K_W]}} = \begin{cases} \frac{1}{v(k,k)}, & k \in K_B, \\ 1, & k \in K_W, \end{cases} \tag{3.2.23}$$

is well defined for any $k \in K$.

(ii) Matrix $(I_W - V)^{-1}$ exists.

Proof. (i) The proof uses the same idea as that of Lemma 3.1.1. Because of Assumption 3.2.10 (i) and (ii) there must be a positive rate $v(k,m) > 0$ to leave any blocking state $k \in K_B$. The generator property $|v(k,k)| = \sum_{h\neq k} v(k,h)$ of the matrix V proves the inequality (3.2.22).

(ii) Because of Assumption 3.2.10 (i) and (ii) the system will leave any state $k \in K_B$ to enter some state $m \in K_W$ in a finite number of jumps caused by generator V. Therefore the matrix $(I_W - V)$ is an essentially diagonal dominant matrix in a finite dimensional space $\mathbb{R}^{K\times K}$. It is invertible according to Lemma A.1.3. □

We now define similar to (3.1.11) in Lemma 3.1.6 a matrix W and determine an explicit representation.

Remark 3.2.13. To simplify notation we will use

$$\{Z(t) = (n,[0,r],k)\} \quad \text{for} \quad \{X(t) = n, R_{\mathrm{S}}(t) \in [0,r], Y(t) = k\}$$

and

$$\{\hat{Z}_0 = (n,0,k)\} \quad \text{for} \quad \{X(0) = n, R_{\mathrm{S}}(0^-) = 0, Y(0) = k\}.$$

Lemma 3.2.14. *For the system from Definition 3.2.9 we define for $k, m \in K$ and $r \in \mathbb{R}_0^+$*

$$W_{km}(r) := P\left(Z(\sigma_1) = (1,[0,r],m) | \hat{Z}_0 = (0,0,k)\right), \tag{3.2.24}$$

and remark that $W_{km}(r) = 0$ for all $m \in K_B$. Then it holds

$$W(r) = G_1(r)(I_W - V)^{-1} I_W. \tag{3.2.25}$$

Proof. Let $k \in K$ and $m \in K_W$. Recall that $R_{\mathrm{S}}(t) = 0$ if $X(t) = 0$ therefore $\{\hat{Z}_0 = (0,0,k)\} \subset \{Z(0) = (0,0,k)\}$. Furthermore by definition $R_{\mathrm{S}}(0^-) = 0$ Therefore

$$\begin{aligned}\{Z(0) = (0,0,k)\} \subset \{(X,Y)(0) = (0,k)\} &= \{(X,Y)(0) = (0,k), R_{\mathrm{S}}(0^-) = 0\} \\ &= \{\hat{Z}_0 = (0,0,k)\}.\end{aligned}$$

Thus we have $\{\hat{Z}_0 = (0,0,k)\} = \{Z_0(0) = (0,0,k)\}$.

Basically, the matrix W has the same structure as W in Proposition 3.1.8, but we will derive a new representation, which is more suitable in the subsequent proofs. Using a similar transformation as in Proposition 3.1.8 we get by a first entrance argument

$$
\begin{aligned}
& W_{km}(r) \\
&= P\left(Z(\sigma_1) = (1,[0,r],m) \middle| \hat{Z}_0 = (0,0,k)\right) \\
&= P\left(Z(\sigma_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right) \\
&= \sum_{h \in K \setminus \{k\}} P\left(Z(\sigma_1) = (1,[0,r],m) \cap Z(\zeta_1) = (0,\mathbb{R}_0^+,h) \middle| Z(0) = (0,0,k)\right) \\
& \quad + \delta_{km} P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right) \\
&= \sum_{h \in K \setminus \{k\}} P\left(Z(\sigma_1) = (1,[0,r],m) \middle| Z(\zeta_1) = (0,\mathbb{R}_0^+,h), Z(0) = (0,0,k)\right) \\
& \quad \cdot P\left(Z(\zeta_1) = (0,\mathbb{R}_0^+,h) \middle| Z(0) = (0,0,k)\right) \\
& \quad + \delta_{km} P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right).
\end{aligned}
$$

Because $R_{\mathrm{S}}(t) = 0$ for $X(t) = 0$ for all t, it holds

$$\{Z(\zeta_1) = (0,\mathbb{R}_0^+,h)\} = \{Z(\zeta_1) = (0,0,h)\}$$

and therefore

$$
\begin{aligned}
& W_{km}(r) \\
&= \sum_{h \in K \setminus \{k\}} P\left(Z(\sigma_1) = (1,[0,r],m) \middle| Z(\zeta_1) = (0,0,h), Z(0) = (0,0,k)\right) \\
& \quad \cdot P\left(Z(\zeta_1) = (0,0,h) \middle| Z(0) = (0,0,k)\right) \\
& \quad + \delta_{km} P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right).
\end{aligned}
$$

Using strong Markov property we get

$$
\begin{aligned}
\iff W_{km}(r) = & \sum_{h \in K \setminus \{k\}} P\left(Z(\sigma_1) = (1,[0,r],m) \middle| Z(0) = (0,0,h)\right) \\
& \cdot P\left(Z(\zeta_1) = (0,0,h) \middle| Z(0) = (0,0,k)\right) \\
& + \delta_{km} P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right)
\end{aligned}
$$

$$
\begin{aligned}
\iff W_{km}(r) = & \sum_{h \in K \setminus \{k\}} \underbrace{P\left(Z(\sigma_1) = (1,[0,r],m) \middle| \hat{Z}_0 = (0,0,h)\right)}_{=W_{hm}(r)} \\
& \cdot P\left(Z(\zeta_1) = (0,0,h) \middle| Z(0) = (0,0,k)\right) \\
& + \delta_{km} P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right).
\end{aligned}
$$

If $k \neq h$ the expression $P\left(Z(\zeta_1) = (0,0,h) \middle| Z(0) = (0,0,k)\right)$ is $\frac{v(k,h)}{-v(k,k)}$ for $k \in K_B$ and 0 for $k \in K_W$. In both cases we will use the expression $\frac{v(k,h)}{-v(k,k)+1_{[k\in K_W]}}$, which is defined for any $k \in K$ (see Lemma 3.2.12 (i))

$$\begin{aligned}\iff W_{km}(r) = &\sum_{h\in K\setminus\{k\}} W_{hm}(r)\frac{v(k,h)}{-v(k,k)+1_{[k\in K_W]}} \\ &+\delta_{km} P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right).\end{aligned}$$

For $k \in K_B$ $P(Z(\zeta_1) = \left(1,[0,r],m) \middle| Z(0) = (0,0,k)\right)$ is zero. For $k \in K_W$ the dynamics of the process is the same as for a queue without environment and the environment states stay unchanged, i.e.

$$P\left(Z(\zeta_1) = (1,[0,r],m) \middle| Z(0) = (0,0,k)\right) = \delta_{km} G_1(r) 1_{[k\in K_W]}$$

with distribution G_1 on page 109.

$$\begin{aligned}\iff \quad & W_{km}(r) \\ = &\sum_{h\in K\setminus\{k\}} W_{hm}(r)\frac{v(k,h)}{-v(k,k)+1_{[k\in K_W]}} + \delta_{km} G_1(r) 1_{[k\in K_W]} \\ = &\sum_{h\in K\setminus\{k\}} \frac{v(k,h)}{-v(k,k)+1_{[k\in K_W]}} W_{hm}(r) + \delta_{km} \underbrace{\frac{1}{-v(k,k)+1_{[k\in K_W]}}}_{1 \text{ for } k\in K_W} 1_{[k\in K_W]} G_1(r).\end{aligned}$$

This equation reads in matrix form

$$W(r) = (-diag(V) + I_W)^{-1}\left((V - diag(V))W(r) + I_W G_1(r)\right)$$

and can finally be transformed into the lemma's statement (3.2.25):

$$\begin{aligned}&(-diag(V) + I_W)W(r) = (V - diag(V))W(r) + I_W G_1(r) \\ &\iff (I_W - V)W(r) = I_W G_1(r) \overset{\text{Lemma 3.2.12 (ii)}}{\Longrightarrow} W(r) = G_1(r)(I_W - V)^{-1} I_W.\end{aligned}$$

□

Definition 3.2.15. For $i > 0$, $n \in \mathbb{N}_0$, $k \in K$ we define

$$U_{km}^{(i,n)}(r) := P\left(\left(X(\tau_1), R_S(\tau_1^-), Y(\tau_1^-)\right) = (i+n-1, 0, m) \middle| Z(0) = (i,r,k)\right), \quad \forall r \in \mathbb{R}^+$$

$$\text{and} \qquad U_{km}^{(i,n)}(0) := 0.$$

Lemma 3.2.16. *For $i > 0$, $n \in \mathbb{N}_0$, and $U_{km}^{(i,n)}(r)$ from Definition 3.2.15 it holds*

$$U_{km}^{(i,n)}(r) = \sum_{h\in K\setminus\{k\}} \frac{v(k,h)}{-v(k,k)+1_{[k\in K_W]}} U_{hm}^{(i,n)}(r) \qquad \forall k \in K_B, r \in \mathbb{R}_0^+. \tag{3.2.26}$$

Proof. For $r = 0$ equation (3.2.26) is obvious. For $r \in \mathbb{R}^+$ and for all $n \in \mathbb{N}_0$ we have $R(\tau_n^-) = 0$ and $U_{km}^{(i,n)}(r)$ can be simplified to

$$U_{km}^{(i,n)}(r) = P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, m) \middle| Z(0) = (i, r, k)\right).$$

Let $k \in K_B$ and $r \in \mathbb{R}^+$. Because $\zeta_1 < \sigma_1$ and $\zeta_1 < \tau_1$ when $k \in K_B$ we have

$$\begin{aligned}
&U_{km}^{(i,n)}(r) \\
&= P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, m) \middle| Z(0) = (i, r, k)\right) \\
&= \sum_{h \in K \setminus \{k\}} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, m) \cap Z(\zeta_1) = (i, r, h) \middle| Z(0) = (i, r, k)\right) \\
&= \sum_{h \in K \setminus \{k\}} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, m) \middle| Z(\zeta_1) = (i, r, h) \cap Z(0) = (i, r, k)\right) \\
&\quad \cdot P\left(Z(\zeta_1) = (i, r, h) \middle| Z(0) = (i, r, k)\right) \\
&\overset{SM}{=} \sum_{h \in K \setminus \{k\}} P\left(\left(X(\tau_1), Y(\tau_1^-)\right) = (i+n-1, m) \middle| Z(0) = (i, r, h)\right) \\
&\quad \cdot P\left(Z(\zeta_1) = (i, r, h) \middle| Z(0) = (i, r, k)\right) \\
&= \sum_{h \in K \setminus \{k\}} U_{hm}^{(i,n)}(r) \cdot P\left(Z(\zeta_1) = (i, r, h) \middle| Z(0) = (i, r, k)\right).
\end{aligned}$$

We analyze the expression $P\left(Z(\zeta_1) = (i, r, h) | Z(0) = (i, r, k)\right)$: By the model assumption it is $\frac{v(k,h)}{-v(k,k)}$ for $k \in K_B$ and zero for $k \in K_W$, therefore we can write it as $\frac{v(k,h)}{-v(k,k)+1_{[k \in K_W]}} 1_{[k \in K_B]}$.

$$\Longleftrightarrow U_{km}^{(i,n)}(r) = \sum_{h \in K \setminus \{k\}} U_{hm}^{(i,n)}(r) \frac{v(k,h)}{-v(k,k) + 1_{[k \in K_W]}}, \qquad \forall k \in K_B, r \in \mathbb{R}^+.$$

$\square$

Lemma 3.2.17. *For $i > 0$, $n \in \mathbb{N}_0$, and $U_{km}^{(i,n)}(r)$ from Definition 3.2.15 it holds*

$$U_{km}^{(i,n)}(r) = \frac{1}{-v(k,k) + 1_{[k \in K_W]}} \delta_{km} \tilde{p}(i, n, r), \qquad \forall k \in K_W, r \in \mathbb{R}_0^+, \tag{3.2.27}$$

with $\tilde{p}(i, n, r)$ from (3.2.1).

Proof. For $r = 0$ equation (3.2.27) follows immediately from definition of $U_{km}^{(i,n)}(0) = 0$ and $\tilde{p}(i, n, 0) = 0$. Let $r \in \mathbb{R}^+$. The factor $\frac{1}{-v(k,k)+1_{[k \in K_W]}}$ is 1 for $k \in K_W$ by (3.2.23).

The factor δ_{km} follows from the model assumption that the environment cannot change when the system is not blocked.

When the system is not blocked, its dynamics is by Definition 3.2.9 (1) the same as, that of a queue without environment. This is determined by the transition probability $\tilde{p}(i, n, r)$. $\square$

Lemma 3.2.18. *In the system from Definition 3.2.9, for any $i > 0$, $n \in \mathbb{N}_0$, and $U^{(i,n)}_{km}(r)$ from Definition 3.2.15 it holds*

$$U^{(i,n)}(r) = \tilde{p}(i,n,r)(I_W - V)^{-1} I_W. \tag{3.2.28}$$

Proof. From Lemma 3.2.16 and Lemma 3.2.17 it follows any $n \geq 0$, $i > 0$ and $r > 0$

$$\begin{aligned} &U^{(i,n)}_{km}(r) \\ &= \sum_{h \in K \setminus \{k\}} \frac{v(k,h)}{-v(k,k) + 1_{[k \in K_W]}} U^{(i,n)}_{hm}(r) 1_{[k \in K_B]} + \frac{1}{-v(k,k) + 1_{[k \in K_W]}} \delta_{km} \tilde{p}(i,n,r) 1_{[k \in K_W]}. \end{aligned}$$

The equation above, written in matrix form, reads

$$\begin{aligned} U^{(i,n)}(r) &= (-diag(V) + I_W)^{-1} \Big((V - diag(V)) U^{(i,n)}(r) + \tilde{p}(i,n,r) I_W \Big) \\ \iff (-diag(V) + I_W) U^{(i,n)}(r) &= ((V - diag(V)) U^{(i,n)}(r) + \tilde{p}(i,n,r) I_W) \\ \iff (I_W - V) U^{(i,n)}(r) &= \tilde{p}(i,n,r) I_W \\ \iff U^{(i,n)}(r) &= \tilde{p}(i,n,r)(I_W - V)^{-1} I_W. \end{aligned}$$

□

Lemma 3.2.19. *In the system from Definition 3.2.9 we define for any $i > 0$, $n \geq 0$, $k \in K$, and $m \in K$*

$$U^{(i,n)}_{km} := P\left((X(\tau_1), R_S(\tau_1), Y(\tau_1^-)) = (i + n - 1, 0, m) | Z_0 = (i, 0, k) \right).$$

Then it holds

$$U^{(i,n)} = \tilde{p}(i,n)(I_W - V)^{-1} I_W. \tag{3.2.29}$$

Proof. According to the model, as soon as the first customer leaves a system, that is right after time 0^-, the new residual requested time R_S is selected at time instant 0. If at this time instant there are i customers in the system, the new value of R_S is distributed according to distribution G_i. Therefore it holds

$$\begin{aligned} U^{(i,n)}_{km} &= \int_0^\infty U^{(i,n)}_{km}(r) dG_i(r) \\ &= \int_0^\infty \tilde{p}(i,n,r)(I_W - V)^{-1} I_W dG_i(r). \end{aligned}$$

Applying (3.2.4) we get (3.2.29). □

Lemma 3.2.20. *In the system from Definition 3.2.9 we define for any $i > 0$, $n \geq 0$, $k \in K$ and $m \in K$*

$$A^{(i,n)}_{km} := P\Big((X(\tau_1), R_S(\tau_1^-), Y(\tau_1)) = (i + n - 1, 0, m) \Big| \hat{Z}_0 = (i, 0, k) \Big).$$

Then holds

$$A^{(i,n)} = U^{(i,n)} R = \tilde{p}(i,n)(I_W - V)^{-1} I_W R. \tag{3.2.30}$$

Proof. In a completely similar way as in Lemma 3.1.6 we can show the following representations

$$\begin{aligned}
A_{km}^{(i,n)} &= P\left(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\right) = (i+n-1, 0, m) \middle| \hat{Z}_0 = (i, 0, k)\right) \\
&= \sum_{h \in K} P\left(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1^-)\right) = (i+n-1, 0, h) \middle| \hat{Z}_0 = (i, 0, k)\right) R(h, m) \\
&= \sum_{h \in K} U_{kh}^{(i,n)} R(h, m).
\end{aligned}$$

□

Lemma 3.2.21. *In the system from Definition 3.2.9 we define for any $i > 0$, $n \geq 0$, $k \in K$, $m \in K$ and $r > 0$*

$$\begin{aligned}
A_{km}^{(i,n)}(r) := P\Big(\left(X(\tau_1), R_S(\tau_1^-), Y(\tau_1)\right) &= (i+n-1, 0, m) \\
&\Big| Z(0) = (i, r, k)\Big).
\end{aligned}$$

Then it holds

$$A^{(i,n)}(r) = U^{(i,n)}(r) R = \tilde{p}(i, n, r)(I_W - V)^{-1} I_W R. \tag{3.2.31}$$

Proof. In a similar way as in Lemma 3.2.20 we show

$$\begin{aligned}
A_{km}^{(i,n)}(r) &= P\left(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\right) = (i+n-1, 0, m) \middle| Z(0) = (i, r, k)\right) \\
&= \sum_{h \in K} P\left(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1^-)\right) = (i+n-1, 0, h) \middle| Z(0) = (i, r, k)\right) R(h, m) \\
&= \sum_{h \in K} U_{kh}^{(i,n)}(r) R(h, m).
\end{aligned}$$

□

Lemma 3.2.22. *In the system from Definition 3.2.9 we define for any $n \geq 0$, $k \in K$, and $m \in K$*

$$B_{km}^{(n)} := P\Big(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\right) = (n, 0, m) \Big| \hat{Z}_0 = (0, 0, k)\Big).$$

Then it holds

$$B^{(n)} = \left((I_W - V)^{-1} I_W\right) A^{(1,n)} = p(1, n) \left((I_W - V)^{-1} I_W\right)^2 R. \tag{3.2.32}$$

Proof.

$$\begin{aligned}
B_{km}^{(n)} &= P\Big(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\right) = (n, 0, m) \Big| \hat{Z}_0 = (0, 0, k)\Big) \\
&= \int_0^\infty \sum_{h \in K} P\Big(\left(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\right) = (n, 0, m) \\
&\qquad\qquad \Big| \hat{Z}_0 = (0, 0, k) \cap \{R_{\mathrm{S}}(\sigma_1) = r, Y(\sigma_1) = h\}\Big) \\
&\qquad dP\left(R_{\mathrm{S}}(\sigma_1) \in [0, r], Y(\sigma_1) = h \middle| \hat{Z}_0 = (0, 0, k)\right).
\end{aligned}$$

Recall that $X(\sigma_1) = 1$ by definition of the stopping time σ_1 if the system is stared empty.

$$\begin{aligned}
\iff \quad & B_{km}^{(n)} \\
= & \sum_{h\in K} \int_0^\infty P\Big(\big(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\big) = (n,0,m) \\
& \qquad \Big| \hat{Z}_0 = (0,0,k) \cap \{X(\sigma_1) = 1, R_{\mathrm{S}}(\sigma_1) = r, Y(\sigma_1) = h\}\Big) \\
& \qquad \cdot dP\left(X(\sigma_1) = 1, R_{\mathrm{S}}(\sigma_1) \in [0,r], Y(\sigma_1) = h \Big| \hat{Z}_0 = (0,0,k)\right).
\end{aligned}$$

Note, for fixed values k and h the function

$$P\left(X(\sigma_1) = 1, R_{\mathrm{S}}(\sigma_1) \in [0,r], Y(\sigma_1) = h \Big| \hat{Z}_0 = (0,0,k)\right)$$

defines a measure on $\mathbb{B}_0^+$.
Using strong Markov property of Z and that $\sigma_1 > 0$ we get

$$\begin{aligned}
\iff \quad & B_{km}^{(n)} \\
= & \sum_{h\in K} \int_0^\infty P\Big(\big(X(\tau_1), R_{\mathrm{S}}(\tau_1^-), Y(\tau_1)\big) = (n,0,m) \\
& \qquad \Big| X(0) = 1, R_{\mathrm{S}}(0) = r, Y(0) = h\Big) \\
& \qquad \cdot dP\left(X(\sigma_1) = 1, R_{\mathrm{S}}(\sigma_1) \in [0,r], Y(\sigma_1) = h \Big| \hat{Z}_0 = (0,0,k)\right) \\
= & \sum_{h\in K} \int_0^\infty A_{hm}^{(1,n)}(r) dW_{kh}(r).
\end{aligned}$$

Using formula (3.2.31) for $A_{hm}^{(1,n)}(r)$ and (3.2.25) for $W_{kh}(r)$ we get

$$\begin{aligned}
\iff \quad & B_{km}^{(n)} \\
= & \sum_{h\in K} \int_0^\infty \tilde{p}(1,n,r) \left((I_W - V)^{-1} I_W R\right)_{hm} d\left(G_1(r)\left((I_W - V)^{-1} I_W\right)_{kh}\right) \\
= & \sum_{h\in K} \left((I_W - V)^{-1} I_W\right)_{kh} \left((I_W - V)^{-1} I_W R\right)_{hm} \int_0^\infty \tilde{p}(1,n,r) dG(r).
\end{aligned}$$

Using $\int_0^\infty \tilde{p}(1,n,r) d\left(G_1(r)\right) = \tilde{p}(1,n)$ we can write $B^{(n)}$in matrix form

$$B^{(n)} \quad = \quad \left((I_W - V)^{-1} I_W\right) \underbrace{\tilde{p}(1,n)\left((I_W - V)^{-1} I_W R\right)}_{=A^{(1,n)}}.$$

$\square$

We are now prepared to evaluate the transition matrix of the $M/G/1/\infty$ system in a random environment from Definition 3.2.9. It turns out that it has precisely the form

(3.2.21). After the construction of matrices $B^{(n)}$ in (3.2.32) and $A^{(i,n)}$ in (3.2.30) we collected enough information for the description of the Markov chain $\left(\hat{Z}_n : n \in \mathbb{N}_0\right)$. In particular we do not need the value $\hat{R}_{\mathrm{S}}(n) = R_{\mathrm{S}}(\tau_n^-)$ which is constant 0 for all n. Starting from here, we focus only on the Markov chain $\left((\hat{X}_n, \hat{Y}_n) : n \in \mathbb{N}_0\right)$.

Lemma 3.2.23. *Consider the continuous time Markov state process of the system described in Definition 3.2.9, and the Markov chain* $(\hat{X}, \hat{Y})$*, embedded at departure instants of customers.*

The one-step transition matrix $\mathbf{P}$ *of* $(\hat{X}, \hat{Y})$

$$\mathbf{P} = \begin{pmatrix} B^{(0)} & B^{(1)} & B^{(2)} & B^{(3)} & \dots \\ A^{(0,1)} & A^{(1,1)} & A^{(2,1)} & A^{(3,1)} & \dots \\ 0 & A^{(0,2)} & A^{(1,2)} & A^{(2,2)} & \dots \\ 0 & 0 & A^{(0,3)} & A^{(1,3)} & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix},$$

is build up by the following block matrices:

$$A^{(i,n)} = \tilde{p}(i,n)H, \qquad \forall i > 0, n \in \mathbb{N}_0, \tag{3.2.33}$$

and

$$B^{(n)} = \tilde{p}(1,n)H, \qquad \forall n \in \mathbb{N}_0, \tag{3.2.34}$$

with

$$H := (I_W - V)^{-1} I_W R. \tag{3.2.35}$$

Proof. Equation (3.2.33) follows from (3.2.30). For (3.2.34) we need to analyze matrix $\left((I_W - V)^{-1} I_W\right)$. It has the block structure

$$(I_W - V) = \left(\begin{array}{c|cc} & K_W & K_B \\ \hline K_W & I & 0 \\ K_B & -V|_{K_B \times K_W} & -V|_{K_B \times K_B} \end{array}\right)$$

$$\Longrightarrow (I_W - V)^{-1} = \left(\begin{array}{c|cc} & K_W & K_B \\ \hline K_W & I & 0 \\ K_B & (I_W - V)^{-1}|_{K_B \times K_W} & (I_W - V)^{-1}|_{K_B \times K_B} \end{array}\right)$$

$$\Longrightarrow (I_W - V)^{-1} I_W = \left(\begin{array}{c|cc} & K_W & K_B \\ \hline K_W & I & 0 \\ K_B & (I_W - V)^{-1}|_{K_B \times K_W} & 0 \end{array}\right).$$

By direct evaluation, this leads to the useful property

$$(I_W - V)^{-1} I_W (I_W - V)^{-1} I_W = (I_W - V)^{-1} I_W. \tag{3.2.36}$$

Which simplifies (3.2.32) to

$$B^{(n)} = p(1,n) \left((I_W - V)^{-1} I_W\right) R = p(1,n) H.$$

□

The next step is similar to that in case of the purely exponential system.

Lemma 3.2.24. *The matrix $H = (I_W - V)^{-1} I_W R$ defined in* (3.2.35) *is a stochastic matrix and there exists a stochastic solution $\hat{\theta}$ of the steady-state equation*

$$\hat{\theta} H = \hat{\theta}. \tag{3.2.37}$$

Proof. The generator property of V leads to

$$(I_W - V)\mathbf{e} = I_W \mathbf{e} + \underbrace{V\mathbf{e}}_{=0} = I_W \mathbf{e} \tag{3.2.38}$$

and the stochasticity of R yields $R\mathbf{e} = \mathbf{e}$. Inserting this into the definition of H leads to

$$H\mathbf{e} = (I_W - V)^{-1} I_W R\mathbf{e} = (I_W - V)^{-1} I_W \mathbf{e} \overset{(3.2.38)}{=} (I_W - V)^{-1}(I_W - V)\mathbf{e} = \mathbf{e}.$$

Since the matrix $\tilde{p}(i,n)(I_W - V)^{-1} I_W$ describes transition probabilities, all its entries are non-negative, therefore the matrix H is stochastic.

Finally, finiteness of K guarantees the existence of a stochastic solution of (3.2.37). $\square$

Theorem 3.2.25. *Consider the $M/G/1/\infty$ in a random environment from Definition 3.2.9 with state dependent service speeds and state dependent selection of requested service times. For the Markov chain $(\hat{X}, \hat{Y})$ embedded at departure points of customers the unique steady-state distribution denoted by $\hat{\pi}$ exists.*

Then $\hat{\pi}$ has product form according to

$$\hat{\pi}(n,k) = \hat{\xi}(n)\hat{\theta}(k), \quad (n,k) \in \mathbb{N}_0 \times K.$$

Here $\hat{\xi}$ is the steady-state distribution of the Markov chain with one-step transition matrix (3.2.2) *derived for the queue length process at departure points in a system with the same parameter as under consideration but without environment, that is a solution of*

$$\hat{\xi}\tilde{P} = \hat{\xi}, \tag{3.2.39}$$

and $\hat{\theta}$ is a stochastic solution of the equation

$$\hat{\theta} H = \hat{\theta}(I_W - V)^{-1} I_W R = \hat{\theta}. \tag{3.2.40}$$

Proof. According to Lemma 3.2.23 the transition matrix $\mathbf{P}$ of the system has block form (3.2.21), which is the tensor product of $\tilde{P}$ from (3.2.2) and H:

$$\mathbf{P} = \tilde{P} \otimes H.$$

Let $\hat{\xi}$ be the steady-state solution of (3.2.39), i.e., of the pure queuing system without environment.

Let $\hat{\theta}$ be the stochastic solution of the equation $\hat{\theta} H = \hat{\theta}$, which exists according to Lemma 3.2.24. Then from tensor calculus of matrices [Neu81, (2.2.19) on p. 53] $\hat{\pi}(n,k) = \hat{\xi}(n)\hat{\theta}(k)$ solves the steady-state equation

$$\hat{\pi}\mathbf{P} = (\hat{\xi} \otimes \hat{\theta})\tilde{P} \otimes H = (\hat{\xi}\tilde{P}) \otimes (\hat{\theta} H) = \hat{\xi} \otimes \hat{\theta} = \hat{\pi}.$$

$\square$

3.3. Applications

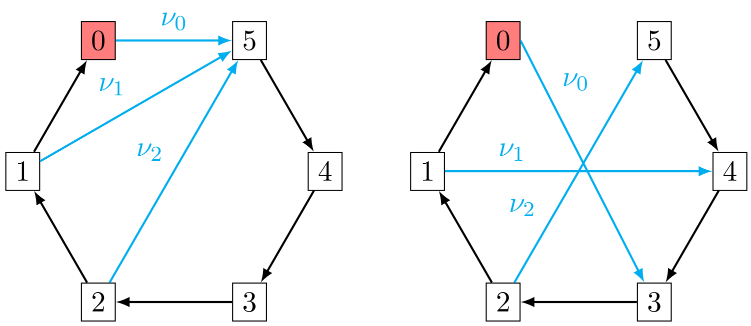
We apply the results from Sections 3.1.1, 3.1.2, and 3.2 to queueing-inventory systems which are dealt with in literature recently, see the review in [KLM11].

Note that to obtain the steady-state distribution $\hat{\pi}$ from π using (3.1.41) could be easier than using (3.1.42), since starting from (3.1.41) does not require to calculate the inverse of $(\lambda I_W - V)$. Nevertheless, we will use (3.1.42) and calculate $WR = \lambda(\lambda I_W - V)^{-1} I_W R$ explicitly to gain more insight into the mathematical structure of the problem.

In any of the following applications the queueing system represents a production facility where raw material arrives and to assemble a final product from a piece of raw material exactly one item from the stock is needed. This item will formally be taken from the stock when the production of the final product is finished.

In Proposition 3.3.1 and Proposition 3.3.2 we slightly extend the lost sales problems from Example 2.2.2 and Example 2.2.3 by incorporating stock size dependent delivery rate ν_k. The main reason of this modification is besides of having more versatile models to demonstrate how each entry of the matrices V and R influences the transition probabilities $\lambda(\lambda I_W - V)^{-1} I_W R$ and the steady-state distribution $\hat{\theta}$.

3.3.1. Systems with exponential service requests



(a) $M/M/1/\infty$ system with $(r = 2, S = 5)$-policy and lost sales.

(b) $M/M/1/\infty$ system with $(r = 2, Q = 3)$-policy and lost sales.

Figure 3.3.1.: Environment transition and interaction diagram for the environment of the lost sales of Proposition 3.3.1 and Proposition 3.3.2. The environment counts the stock size of the inventory.

Proposition 3.3.1. *We consider an exponential single server queue with state dependent service rates, environment dependent replenishment rates, and an attached inventory under (r, S) policy (with $0 \leq r < S \in \mathbb{N}$), and lost sales when the inventory is depleted.*

Using the definitions of Section 3.1.1 we set the environment state space $K := \{0, \ldots, S\}$ with $K_B = \{0\}$, $X(t)$ is the queue length at time t, and $Y(t) = k$ indicates that at time t

the stock contains exactly k items. The strictly positive transitions intensities are

$$\begin{aligned} q((n,k),(n+1,k)) &= \lambda, && k > 0,\\ q((n,k),(n,S)) &= \nu_k, && 0 \le k \le r,\\ q((n,k),(n-1,k-1)) &= \mu(n), && n > 0, 1 \le k \le S,\\ q((n,k),(l,m)) &= 0, && \textit{otherwise.} \end{aligned}$$

The steady-state $\hat{\pi}$ of the Markov chain $(\hat{X}, \hat{Y})$ embedded at departure times has product form

$$\hat{\pi}(n,k) = \xi(n)\hat{\theta}(k), \quad (n,k) \in \mathbb{N}_0 \times K, \tag{3.3.1}$$

with

$$\xi(n) := C^{-1}\left(\prod_{i=1}^{n} \frac{\lambda}{\mu(i)}\right), \quad n \in \mathbb{N}_0,$$

with normalization constant

$$C := \sum_{n=0}^{\infty}\left(\prod_{i=1}^{n} \frac{\lambda}{\mu(i)}\right)$$

and

$$\hat{\theta}(k) = \begin{cases} C_2^{-1} \cdot \prod_{i=1}^{k} \left(\frac{\lambda+\nu_i}{\lambda}\right)^i, & 0 \le k \le r,\\ C_2^{-1} \cdot \prod_{i=1}^{r} \left(\frac{\lambda+\nu_i}{\lambda}\right)^i, & r+1 \le k \le S-1,\\ 0, & k = S, \end{cases} \tag{3.3.2}$$

with

$$C_2 = \sum_{k=0}^{r-1}\prod_{i=1}^{k}\left(\frac{\lambda+\nu_i}{\lambda}\right)^i + (S-r)\prod_{i=1}^{r}\left(\frac{\lambda+\nu_i}{\lambda}\right)^i.$$

Note that even for the constant values $\nu_k = \nu$ the marginal distribution $\hat{\theta}$ in (3.3.2) differs from the marginal stationary distribution $P(Y(t) = k)$ of the continuous time process in Example 2.2.2.

Proof. According to Corollary 3.1.17 the marginal distribution $\hat{\theta}$ is the stochastic solution of the equation

$$\hat{\theta}\,\underbrace{\lambda(\lambda I_W - V)^{-1} I_W}_{=W}\, R = \hat{\theta}. \tag{3.3.3}$$

We calculate the matrix $\lambda(\lambda I_W - V)^{-1} I_W R$ explicitly.

$$(\lambda I_W - V) =$$

$$\left(\begin{array}{c|cccccccccc} & 0 & 1 & 2 & \dots & r-1 & r & r+1 & \dots & S-1 & S\\ \hline 0 & \nu_0 & 0 & 0 & & 0 & 0 & 0 & & 0 & -\nu_0\\ 1 & 0 & (\nu_1+\lambda) & 0 & & 0 & 0 & 0 & & 0 & -\nu_1\\ 2 & 0 & 0 & (\nu_2+\lambda) & & 0 & 0 & 0 & & 0 & -\nu_2\\ \vdots & \vdots & \vdots & \ddots & & & & & & & \vdots\\ r & 0 & 0 & 0 & \dots & 0 & (\nu_r+\lambda) & 0 & \dots & 0 & -\nu_r\\ r+1 & 0 & 0 & 0 & & 0 & 0 & \lambda & \dots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots\\ S-1 & 0 & 0 & 0 & \dots & 0 & 0 & & & \lambda & 0\\ S & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 0 & \lambda \end{array}\right),$$

$$W = \lambda(\lambda I_W - V)^{-1} I_W =$$

$$\left(\begin{array}{c|cccccccccc}
 & 0 & 1 & 2 & \dots & r-1 & r & r+1 & \dots & S-1 & S \\ \hline
0 & 0 & 0 & 0 & & 0 & 0 & 0 & & 0 & 1 \\
1 & 0 & \frac{\lambda}{\nu_1+\lambda} & 0 & & 0 & 0 & 0 & & 0 & \frac{\nu_1}{\nu_1+\lambda} \\
2 & 0 & 0 & \frac{\lambda}{\nu_2+\lambda} & & 0 & 0 & 0 & & 0 & \frac{\nu_2}{\nu_2+\lambda} \\
\vdots & \vdots & \vdots & \ddots & & & & & & & \vdots \\
r & 0 & 0 & 0 & \dots & 0 & \frac{\lambda}{\nu_r+\lambda} & 0 & \dots & 0 & \frac{\nu_r}{\nu_r+\lambda} \\
r+1 & 0 & 0 & 0 & & 0 & 0 & 1 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
S-1 & 0 & 0 & 0 & \dots & 0 & 0 & & & 1 & 0 \\
S & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 0 & 1
\end{array}\right),$$

$$R =$$

$$\left(\begin{array}{c|cccccc}
 & 0 & 1 & 2 & \dots & S-1 & S \\ \hline
0 & 1 & 0 & 0 & \dots & 0 & 0 \\
1 & 1 & 0 & 0 & \dots & 0 & 0 \\
2 & 0 & 1 & 0 & & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
S-1 & 0 & 0 & 0 & \dots & 0 & 0 \\
S & 0 & 0 & 0 & \dots & 1 & 0
\end{array}\right),$$

$$WR = \lambda(\lambda I_W - V)^{-1} I_W R =$$

$$\left(\begin{array}{c|cccccccccc}
 & 0 & 1 & 2 & \dots & r-1 & r & r+1 & \dots & S-1 & S \\ \hline
0 & 0 & 0 & 0 & & 0 & 0 & 0 & & 1 & 0 \\
1 & \frac{\lambda}{\nu_1+\lambda} & 0 & 0 & & 0 & 0 & 0 & & \frac{\nu_1}{\nu_1+\lambda} & 0 \\
2 & 0 & \frac{\lambda}{\nu_2+\lambda}0 & 0 & & 0 & 0 & 0 & & \frac{\nu_2}{\nu_2+\lambda} & 0 \\
\vdots & \vdots & \vdots & \ddots & & & & & & \vdots & 0 \\
r & 0 & 0 & 0 & \dots & \frac{\lambda}{\nu_r+\lambda} & 0 & 0 & & \frac{\nu_r}{\nu_r+\lambda} & 0 \\
r+1 & 0 & 0 & 0 & & 0 & 1 & 0 & & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & 0 \\
S-1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \ddots & 0 & 0 \\
S & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 1 & 0
\end{array}\right).$$

Inserting this and (3.3.2) into (3.3.3) finishes the proof. □

Proposition 3.3.2. *We consider an exponential single server queue with state dependent service rates, environment dependent replenishment rates, and an attached inventory under (r, Q) policy (with $0 \le r < Q \in \mathbb{N}$), and lost sales when the inventory is depleted.*

Using the definitions of Section 3.1.1 we set the environment state space $K := \{0, \dots, S\}$ with $K_B = \{0\}$, $X(t)$ is the queue length at time t, and $Y(t) = k$ indicates that at time t the stock contains exactly k items. The strictly positive transition intensities are

$$\begin{aligned}
q((n,k),(n+1,k)) &= \lambda, && k > 0, \\
q((n,k),(n,k+Q)) &= \nu_k, && 0 \le k \le r, \\
q((n,k),(n-1,k-1)) &= \mu(n), && n > 0, 1 \le k \le r+Q, \\
q((n,k),(l,m)) &= 0, && \textit{otherwise.}
\end{aligned}$$

The steady-state $\hat{\pi}$ has product form

$$\hat{\pi}(n,k) = \xi(n)\hat{\theta}(k), \quad (n,k) \in \mathbb{N}_0 \times K, \tag{3.3.4}$$

with

$$\xi(n) := C^{-1}\left(\prod_{i=1}^{n} \frac{\lambda}{\mu(i)}\right), \qquad n \in \mathbb{N}_0,$$

with normalization constant

$$C := \sum_{n=0}^{\infty}\left(\prod_{i=1}^{n} \frac{\lambda}{\mu(i)}\right)$$

and

$$\hat{\theta}(k) = \begin{cases} C_2^{-1} \cdot \prod_{i=1}^{k} \left(\frac{\lambda+\nu_i}{\lambda}\right)^i, & 0 \le k \le r, \\ C_2^{-1} \cdot \prod_{i=1}^{r} \left(\frac{\lambda+\nu_i}{\lambda}\right)^i, & r+1 \le k \le Q-1, \\ C_2^{-1} \prod_{i=1}^{r} \left(\frac{\lambda+\nu_i}{\lambda}\right)^i - \prod_{i=1}^{k-Q} \left(\frac{\lambda+\nu_i}{\lambda}\right)^i, & Q \le k \le r+Q-1, \\ 0, & k = r+Q, \end{cases} \tag{3.3.5}$$

with normalization constant

$$C_2 = (Q-r)\prod_{i=1}^{k}\left(\frac{\lambda+\nu_i}{\lambda}\right)^i.$$

Proof. According to Corollary 3.1.17 the marginal distribution $\hat{\theta}$ is the stochastic solution of the equation

$$\hat{\theta}\,\underbrace{\lambda(\lambda I_W - V)^{-1} I_W}_{=W} R = \hat{\theta}. \tag{3.3.6}$$

We calculate the matrix $\lambda(\lambda I_W - V)^{-1} I_W R$ explicitly (the remark from Proposition 3.3.1 on indexing the matrices applies here as well).

$$(\lambda I_W - V) =$$

$$\left(\begin{array}{c|ccccccccccccc}
 & 0 & 1 & 2 & \dots & r-1 & r & r+1 & \dots & Q & 1+Q & 2+Q & \dots & r+Q \\ \hline
0 & \nu_0 & 0 & 0 & & 0 & 0 & 0 & & -\nu_0 & & & & 0 \\
1 & 0 & (\nu_1+\lambda) & 0 & & 0 & 0 & 0 & & 0 & -\nu_1 & & & 0 \\
2 & 0 & 0 & (\nu_2+\lambda) & & 0 & 0 & 0 & & 0 & & -\nu_2 & & 0 \\
\vdots & \vdots & \vdots & \ddots & & & & & & & & & \ddots & \vdots \\
r & 0 & 0 & 0 & \dots & 0 & (\nu_r+\lambda) & 0 & \dots & 0 & & & & -\nu_r \\
r+1 & 0 & 0 & 0 & & 0 & 0 & \lambda & \dots & 0 & & & & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & & & & \vdots \\
Q & 0 & 0 & 0 & \dots & 0 & 0 & & & \lambda & & & & 0 \\
1+Q & 0 & 0 & 0 & & 0 & 0 & & & & & & & \\
2+Q & & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & & \\
r+Q & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 0 & & & & \lambda
\end{array}\right),$$

$$W = \lambda(\lambda I_W - V)^{-1} I_W =$$

$$\left(\begin{array}{c|cccccccccccccc}
 & 0 & 1 & 2 & \dots & r-1 & r & r+1 & \dots & Q & 1+Q & 2+Q & \dots & r+Q \\
\hline
0 & 0 & 0 & 0 & & 0 & 0 & 0 & & 1 & 0 & 0 & & 0 \\
1 & 0 & \frac{\lambda}{\nu_1+\lambda} & 0 & & 0 & 0 & 0 & & 0 & \frac{\nu_1}{\nu_1+\lambda} & & & \\
2 & 0 & 0 & \frac{\lambda}{\nu_2+\lambda} & & 0 & 0 & 0 & & 0 & & \frac{\nu_2}{\nu_2+\lambda} & & \\
\vdots & \vdots & \vdots & \ddots & & & & & & & & & \ddots & \\
r & 0 & 0 & 0 & \dots & 0 & \frac{\lambda}{\nu_r+\lambda} & 0 & \dots & 0 & & & & \frac{\nu_r}{\nu_r+\lambda} \\
r+1 & 0 & 0 & 0 & & 0 & 0 & 1 & \dots & 0 & & & & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & & & & \vdots \\
Q & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 1 & & & & 0 \\
1+Q & 0 & 0 & 0 & & 0 & 0 & 0 & & & 1 & & & \\
2+Q & 0 & 0 & 0 & & 0 & 0 & 0 & & & & 1 & & \\
 & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & & & & \ddots & \\
S & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 0 & 0 & 0 & & 1
\end{array}\right),$$

$$R =$$

$$\left(\begin{array}{c|cccccc}
 & 0 & 1 & 2 & \dots & r-1+Q & r+Q \\
\hline
0 & 1 & 0 & 0 & & 0 & 0 \\
1 & 1 & 0 & 0 & & 0 & 0 \\
2 & 0 & 1 & 0 & & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\
r-1+Q & 0 & 0 & 0 & \dots & 0 & 0 \\
r+Q & 0 & 0 & 0 & \dots & 1 & 0
\end{array}\right),$$

$$WR = \lambda(\lambda I_W - V)^{-1} I_W R =$$

$$\left(\begin{array}{c|cccccccccccccc}
 & 0 & 1 & 2 & \dots & r-1 & r & r+1 & \dots & Q-1 & Q & 1+Q & 2+Q & \dots & r+Q \\
\hline
0 & 0 & 0 & 0 & & 0 & 0 & 0 & & 1 & 0 & 0 & 0 & & 0 \\
1 & \frac{\lambda}{\nu_1+\lambda} & 0 & 0 & & 0 & 0 & 0 & & & \frac{\nu_1}{\nu_1+\lambda} & & & & \\
2 & 0 & \frac{\lambda}{\nu_2+\lambda} & & & 0 & 0 & 0 & & & & \frac{\nu_2}{\nu_2+\lambda} & & & \\
\vdots & \vdots & \vdots & \ddots & & & & & & & & & \ddots & & \\
r & 0 & 0 & 0 & \dots & \frac{\lambda}{\nu_r+\lambda} & 0 & 0 & & & & & & \frac{\nu_r}{\nu_r+\lambda} & \\
r+1 & 0 & 0 & 0 & & 0 & 1 & & & & & & & & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & & & & & & \vdots \\
Q-1 & 0 & 0 & 0 & 0 & 0 & 0 & & \ddots & & & & & & \\
Q & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & 1 & & & & & 0 \\
1+Q & 0 & 0 & 0 & & 0 & 0 & 0 & & & 1 & & & & \\
2+Q & 0 & 0 & 0 & & 0 & 0 & 0 & & & & 1 & & & \\
 & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & & & & \ddots & & \\
r+Q & 0 & 0 & 0 & \dots & 0 & 0 & 0 & & & 0 & 0 & 0 & 1 & 0
\end{array}\right).$$

Note that even for constant values $\nu_k = \nu$ the marginal distribution $\hat{\theta}$ under (r, Q) policy differs from the marginal steady-state distribution $P(Y(t) = k)$ in continuous time from Example 2.2.3.

Inserting this and (3.3.5) into (3.3.6) finishes the proof. $\square$

3.3.2. Systems with non-exponential service requests

Proposition 3.3.3. *We consider a single server queue of $M/G/1/\infty$-type, with state dependent service speeds, state dependent selection of requested service times, exponential-ν replenishment times, and an attached inventory under $(r = 0, S)$ policy (with $0 < S \in \mathbb{N}$), and lost sales when the inventory is depleted (see Definition 3.2.9) such that Assumption 3.2.10 holds.*

We have $K = \{S, S-1, \dots, 1, 0\}$ with $K_B = \{0\}$.

The stochastic jump matrix R represents the downward jumps of the inventory

$$R = \left(\begin{array}{c|cc} & 0 \dots S-1 & S \\ \hline 0 & (1, 0, \dots, 0) & 0 \\ 1 & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} & 0 \\ \vdots & & \\ S & & \end{array}\right),$$

and because the environment moves independently only if there is stockout, the environment generator V has only non zero entries $v(0, S) = \nu$, $v(0, 0) = -\nu$. So with $K_B = \{0\}$ the requirement of Theorem 3.2.25 is fulfilled.

$$V = \left(\begin{array}{c|cc} & 0 & 1 \dots S \\ \hline 0 & -\nu & (0, \dots, 0, \nu) \\ 1 & & \\ \vdots & & 0 \\ S & & \end{array}\right).$$

From Theorem 3.2.25 we conclude that the Markov chain $(\hat{X}, \hat{Y})$, embedded at departure instants of customers has a stationary distribution $\hat{\pi}$ of product form

$$\hat{\pi}(n, k) = \hat{\xi}(n)\hat{\theta}(k), \quad (n, k) \in \mathbb{N}_0 \times K.$$

Here $\hat{\xi}$ is the steady-state distribution of the Markov chain with one-step transition matrix (3.2.2) *derived for the queue length process at departure points in a system with the same parameters as under consideration but without environment, i.e., a solution of $\hat{\xi}\tilde{P} = \hat{\xi}$, and $\hat{\theta}$ is for $k \in \{0, 1, \dots, S\}$*

$$\hat{\theta}(k) = \frac{1}{S} \quad k \neq S, \qquad \hat{\theta}(S) = 0. \tag{3.3.7}$$

According to Theorem 3.2.25, $\hat{\theta}$ is a stochastic solution of the equation $\hat{\theta}(I_W - V)^{-1} I_W R = \hat{\theta}$. We calculate the matrix $H = (I_W - V)^{-1} I_W R$ explicitly.

$$(I_W - V) = \left(\begin{array}{c|cc} & 0 & 1 \dots S \\ \hline 0 & \nu & (0, \dots, 0, -\nu) \\ 1 & 0 & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \\ \vdots & \vdots & \\ S & 0 & \end{array}\right)$$

$$(I_W - V)^{-1} = \left(\begin{array}{c|cc} & 0 & 1\dots S \\ \hline 0 & \frac{1}{\nu} & (0,\dots,0,1) \\ 1 & 0 & \\ \vdots & \vdots & \begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix} \\ S & 0 & \end{array}\right)$$

$$(I_W - V)^{-1} I_W = \left(\begin{array}{c|cc} & 0 & 1\dots S \\ \hline 0 & 0 & (0,0,\dots,1) \\ 1 & 0 & \\ \vdots & \vdots & \begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix} \\ S & 0 & \end{array}\right)$$

$$H = (I_W - V)^{-1} I_W R = \left(\begin{array}{c|cc} & 0\dots S-1 & S \\ \hline 0 & (0,0,\dots,1) & 0 \\ 1 & & 0 \\ \vdots & \begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix} & \vdots \\ S & & 0 \end{array}\right)$$

$\hat{\theta}$ defined in (3.3.7) is the unique solution of the equation (3.2.40).

Proposition 3.3.4. *We consider a single server queue of $M/G/1/\infty$-type, with state dependent service speeds, state dependent selection of requested service times, inventory management policy (r, Q) or (r, S), and zero lead times (see Definition 3.2.9, and note that lost sales do not occur because of zero lead time).*

In the case of (r, S) policy the inventory size after the first delivery will stay on between $r+1$ and S, therefore for long term behaviour of the system we take in account only environment states $K = \{r+1, r+2, \dots, S\}$. The zero lead time means $V = 0$, $K_B = \emptyset$, and the corresponding R matrix has the form

$$R = \left(\begin{array}{c|cc} & r+1\dots S-1 & S \\ \hline r+1 & (0,0,\dots,0) & 1 \\ r+2 & & \\ \vdots & \begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix} & 0 \\ S & & \end{array}\right).$$

The steady-state distribution has a product form

$$\hat{\pi}(n, k) = \hat{\xi}(n)\hat{\theta}(k), \quad (n, k) \in \mathbb{N}_0 \times K,$$

with

$$\hat{\theta}(k) = \frac{1}{S-r}, \quad k \in K. \tag{3.3.8}$$

Proof. According to Theorem 3.2.25 $\hat{\theta}$ is a stochastic solution of the equation $\hat{\theta}(I_W - V)^{-1} I_W R = \hat{\theta}$ We calculate the matrix $H = (I_W - V)^{-1} I_W R$, which in the case of the model equivalent to

$$\hat{\theta}\underbrace{(I_W - V)^{-1}}_{=I}\underbrace{I_W}_{=I} R = \hat{\theta} \Longleftrightarrow \hat{\theta} R = \hat{\theta}$$

with a unique stochastic solution (3.3.8).

For system under (r, Q) policy with zero lead times (3.3.8) holds as well, the proof is analogous, we just set $S = r + Q$. $\square$

Remark. Similar results for the steady state of queueing-inventory systems with zero lead times (without speeds) were obtained by Vineetha in [Vin08, Theorem 5.2.1] for the case of i.i.d service times.

Part II.

Product-form networks

Parts of Part II mainly focus on Jackson networks in random environment with constant parameters V and R will be first published in Advances in Applied Probability Vol. 48 No 2 © 2016 by Applied Probability Trust.

4. Introduction

Queueing networks with product-form steady-state distribution have found many fields of applications, e.g. production systems, telecommunications, and computer system modeling. The success of this class of models and its relatives stems from the simple structure of the steady-state distribution which provides access to easy performance evaluation procedures.

Starting from the work of Jackson [Jac57] various generalizations have been developed. A branch of research which recently has found interest are queueing networks in a random environment with product-form steady-state distributions.

For single service stations (in isolation) there is a long history with investigations on the behaviour of the stations under external influences, which are subsumed under the term of an environment. Similarly birth-death processes (as generalizations of classical $M/M/1/\infty$ queues) in a random environment are well investigated. Most of this work resulted in complex steady-state distributions, see e.g. [CT81], [Cog80], [Yec73], [KY12], [Fal96]. A branch of related research is concerned with service systems under external influences which cause the service process to break down or decreases availability of servers, see e.g. the early paper [Zol66], and for recent survey [KPC12]. The results in these articles most often lack the elegance of Jackson's product-form steady state and the simplicity of the steady states of birth-death processes.

Exceptions are [SSD$^+$06], [KN13], and [SAH13], where the environment of a production model (queue) is an associated inventory, and [EF98], [SD03a], where the influence of the environment on the queue results in randomly occurring breakdowns of the server, and [KD14], where the environment of a sensor node encompasses the node's neighbors, their status, etc. In these papers on queueing-environment processes it is shown that the two-dimensional steady-state distribution factorizes into the product of the marginal one-dimensional steady-state distributions, shortly: In equilibrium and in the long run the steady-state distribution for the queue and the environment decouples.

Related work is by Yamazaki and Miyazawa [YM95] where the environment of a queue is called the "set of background states" which determines transition rates of the queue and on the other side is influenced by state changes of the queue. Yamazaki and Miyazawa prove a decomposition property (which is in the spirit of decoupling as in Jackson's theorem) **not** for the queue and the environment but for the joint (QUEUEING-ENVIRONMENT PROCESS) and some supplementary variables which are introduced for Markovisation of the process in case of non-exponential service times and non exponential holding times for the background states.

A first approach to find product-form steady-state distributions for Jackson networks in a random environment seemingly was the work of Zhu [Zhu94]. Economou [Eco03, Eco04, Eco05], Balsamo and Marin [BM13], and Tsitsiashvili, Osipova, Koliev, Baum [TOKB02] continued the investigations. The procedure in these papers for a network of single exponential servers is as follows (explained in terms of Zhu's notation).

The key ingredients for node i in *classical Jackson networks* are an external Poisson-λ_i arrival stream, exponential-μ_i service times, and a Markovian routing scheme, which produces an overall arrival rate η_i and a local marginal stationary distribution π_i with

$$\pi_i(n) = \left(1 - \frac{\eta_i}{\mu_i}\right)\left(\frac{\eta_i}{\mu_i}\right)^n, \quad n \in \mathbb{N}_0. \tag{4.0.1}$$

In [Zhu94] these fundamental parameters depend on the state of the external environment: If this is in state k the parameters are $\lambda_i(k), \mu_i(k), \eta_i(k)$, and with the additional **assumption** that the utilizations $\rho_i(k) := \eta_i(k)/\mu_i(k)$ do not depend on k, i.e.

$$\rho_i(k) = \eta_i(k)/\mu_i(k) = \eta_i/\mu_i =: \rho_i, \tag{4.0.2}$$

say, the local stationary distribution is (4.0.1) again. Zhu and his followers do not explain how independence of k for utilizations emerges, i.e. no control mechanism is described which holds $\eta_i(k)/\mu_i(k)$ invariant during changes of the environment.

In [TOKB02] the authors argue for the invariance $\eta_i(k)/\mu_i(k) = \eta_i/\mu_i$ by pointing on the fact that there exist natural control mechanisms of technical and biological systems to react to changes of the environment.

The quest for such rules has on the other side a long history in related fields, especially in the control of communications networks, where the term "rerouting schemes" describes the necessity to react, e.g. to buffer overflow, broken down nodes, (partial) degrading of transmission lines. Examples are described in [SD03a] under the heading of "skipping", "repeated service–random destination", "stalling". The first regime is called "jump-over protocol" by other authors, see [Dij88].

To put it into a more concrete example: These schemes try to mimic by stylized policies in communications networks an exchange of routing tables or rules for dynamic traffic reallocation to paths as reaction to changes of the environment or the network's load situation, for a discussion see [NSB$^+$03] or [GKK95].

In the range of telecommunication networks queueing networks are models for packet oriented transmission. For circuit switched networks (loss networks) there is a rich literature on dynamic routing schemes to adjust routing in accordance with time varying offered traffic and to provide flexibility and robustness to respond to failures or overloads, see [GK90, Section 1] and the references given there, and [SS12].

Our investigation originates from the observation often made in stochastic networks with blocking or with unreliable servers, that it is possible to obtain explicit product form stationary distributions by implementing a clever rerouting regime for customers who find at a node, selected for his next entrance, the buffer full or the node broken down, see e.g. [SD03a] and the literature cited there. A bulk of examples can be found in the book [Dij93] and in [BvD10, Chapt. 1, 9]. These rerouting schemes maintain the utilization as in (4.0.2) for those nodes which are not blocked, resp. not broken down. These results are not covered by the model of Zhu, where the $\eta_i(k)$ are required to be strictly positive.

Similar product form network models for randomized medium access control protocols have found interest recently because of tractability of the model, see [SS12].

Overview of our results.

(I) Our first contribution is that we complement Zhu's results by providing physically meaningful rerouting schemes which maintain in his setting the utilizations (4.0.2).

(II) Our second contribution is to extend Zhu's results in a way that the mentioned results on networks with breakdowns in [SD03a] are covered and generalized.

We start as Zhu [Zhu94] with a Jackson network with locally queue length, n_i, dependent service intensities, which may depend additionally on the actual state k of the network's environment, $\mu_i(n_i, k)$. The external arrival rates are $\lambda_i(k)$, and the overall arrival rates are $\eta_i(k)$. We will construct control schemes which adapt the routing to changes in the parameters, when the environment changes from k to m and the service intensities from $\mu_i(n_i, k)$ to $\mu_i(n_i, m)$. We will prove that under these new control schemes the respective ratios are maintained constant: $\eta_i(k)/\mu_i(n_i, k) = \eta_i/\mu_i(n_i)$ will be independent of k (but not of n_i) as long as $\eta_i(k) > 0$ holds. Our theorems will cover the case $\eta_i(k) = 0$ as well, which is in force e.g., if an unreliable node i is down and therefore does neither serve nor accept new customers.

The most important consequence will be that for Jackson networks in a random environment we obtain a product form stationary distribution, similar to [Zhu94, Theorem 1], but under much more general assumptions.

(III) Our third contribution is an extension of the environment structures found in the mentioned previous literature because in our setting the network process influences the environment process as well. We emphasize that different from the mentioned work in [Zhu94], [Eco05], [BM13], and [TOKB02], our environment process is not Markov for its own because changes in the queue length processes may enforce the environment to immediate changes as well. There will be a two-way interaction between the service systems and the environment in our model.

The modifications of the random walks which we construct encompass the rerouting schemes which are often found in the literature and are called e.g. "jump-over protocol" or "skipping" or "blocking after service and retrial". Therefore, our results generalize some of those in [Zhu94], [Eco05], [BM13], and [TOKB02], and as well some of [SD03a].

(IV) We start in Section 5 with a detailed study of routing chains for the selection of individual customers' itineraries in the network and suitable modifications of these in terms of general Markov chains, resp. random walks. The modification is realized in analogy to principles occurring in Markov Chain Monte Carlo (MCMC) algorithms by attaching to any state of the chain a (state dependent) "acceptance probability" to operate via Bernoulli experiments on the original transition matrix of the chain, which is considered as "candidate-generating matrix" (see [Bre99, Section 7.1]).

The jump is realized, if accepted, and the chain settles down at the selected state for the next time slot. But other than in MCMC algorithms, "not accepted" in our modification means not necessarily for the random walk to stay on at the departure state. Additionally to the standard one we consider the policy that from the selected, but not accepted state the chain tries again to find a next state, now with probabilities from the "candidate-generating matrix" determined by the row of the not accepted state. Thereafter acceptance is tested again, and so on.

We will show that the steady-state distributions of the original chain and the modified chains are intimately connected and that we can express the new steady state easily in terms of the steady state of the old chain and the acceptance probabilities.

(V) Although our research started with a quest for new rerouting schemes for Jackson networks in case of (partial) non-availability of servers, the developed schemes seem to be of interest for their own.

From an abstract point of view our modification of the original chain can be considered as a complicated change of measure for the process distribution, see [AG07, Chapter V, 1c, Example 8], which results in a surprisingly simple explicit change of measure for the stationary distribution.

The modification algorithms which lead to this change of measure can be distinguished according to the property to be local or global (as discussed in [SS12]) with respect to the one-step transition graph of the original chain.

In Section 5.3 we discuss connections of our randomization algorithms to MCMC algorithm and to von Neumann's acceptance-rejection scheme for sampling from a complicated distribution. Furthermore, we compare the different modifications of the random walk with respect to Peskun ordering.

(VI) Our applications of the randomized random walk algorithms to stochastic networks is in the spirit of "Performability" [Mey84, p. 648]: "If computing system performance is degradable, then . . . evaluation must deal simultaneously with aspects of both performance and reliability." A more recent compendium on these topics is [HMRT01]. We start with a standard Jackson network in Section 6 and consider in Section 6.1 the situation where the service capacities of nodes are changed and by some network control an optimized utilization of the nodes should be maintained. This relates our results to problems of optimal design of networks (e.g. n [Whi85], [Sti09]).

When nodes are only degraded, the network controller's policy is to reject a portion of the load offered and to redistribute the admitted load. This is organized by applying our modification algorithms for random walks to the routing chains for the customers. Additionally to (partial) degrading we can handle with our algorithm complete breakdown of nodes, and we even allow to speed up service at nodes, while others are degraded or broken down. In the latter case the controller additionally offers an increased load to the network. In any case the utilization of the nodes which are not completely down is maintained by the control policy.

While Section 6 describes how to transform a Jackson network into another one where service capacities are degraded and/or upgraded and routing is adapted, maintaining the nodes' utilization and therefore the joint product form stationary distribution, we utilize the obtained transformation rules in Section 7 to tackle a dynamic problem: We adapt routing by different algorithms to the impact of a dynamically changing environment. The environment's changes cause the nodes' service capacities to degrade and/or upgrade, even complete breakdown with following repair or partial repair can be handled by our algorithms for route adaptation, which will immediately achieve the equilibrium queue length distributions anew.

The most surprising result is an iterated product form stationary distribution of the system process, which is a multidimensional Markov process recording jointly the environment's status and the joint queue lengths vector for the network process. The product form says that

- the queue length vector and the environment status, and,
- inside of the joint queue lengths vector, the local queue lengths

asymptotically and in equilibrium decouple (are decomposable in the sense of [YM95]). This result is remarkable as we do not require that the environment process is a Markov process of its own, as it is necessary for the results proved in [Zhu94], [Eco05], [BM13], and [TOKB02]. We can handle a two-way interaction between environment which enforces changes of the nodes' service capacity when it changes, and the queueing network process, which triggers immediate jumps of the environment, when a customer departs from the network. Such (more complicated) two-way interaction is investigated in [YM95] and [Eco05] as well, but due to the more complicated structure they loose in their results decoupling of the queue length and the environment status.

Notation and conventions:

- Throughout, the node set of our graphs (networks) are denoted by $\overline{J} := \{1, \ldots, J\}$, and the "extended node set" is $\overline{J}_0 := \{0, 1, \ldots, J\}$, where "0" refers to the external source and sink of the network.
- $\mathbf{e}_j$ is the standard j-th base vector in $\mathbb{N}_0^{\overline{J}}$ if $1 \leq j \leq J$.
- $\mathbf{n} = (n_j : j \in \overline{J})$ usually is the joint queue length vector of some queueing network.
- For any finite index set $\overline{F} = \{0, 1, \ldots, F\}$ and any $\boldsymbol{\alpha} = (\alpha_j : j \in \overline{F})$ we define the matrices $I_{\boldsymbol{\alpha}}$ which is the diagonal matrix indexed by $\overline{F}$ with α_i on its diagonal, i.e.

  $$I_{\boldsymbol{\alpha}} := \begin{pmatrix} \alpha_0 & & & \\ & \alpha_1 & & \\ & & \ddots & \\ & & & \alpha_F \end{pmatrix},$$

  and similarly we define $I_{(1-\boldsymbol{\alpha})}$ indexed by $\overline{F}$ as

  $$I_{(1-\boldsymbol{\alpha})} := \begin{pmatrix} 1-\alpha_0 & & & \\ & 1-\alpha_1 & & \\ & & \ddots & \\ & & & 1-\alpha_F \end{pmatrix}.$$

  Here and elsewhere we agree that empty entries in a matrix are read to be zero.
- For real valued functions $f, g : \overline{F} \to \mathbb{R}$ on a countable set we define $f \bullet g$ as their point wise multiplication i.e. $f \bullet g(i) = f(i) \cdot g(i), i \in \overline{F}$, and the diagonal matrix $I_{f \bullet g}$ by $I_{f \bullet g}(i, j) := 1_{[i=j]} \cdot f(i) \cdot g(i),\ i, j \in \overline{F}$.
- For $\mathbf{x} = (x_j : j \in \overline{F})$ we define $||\mathbf{x}||_\infty := \sup_{j \in \overline{F}} |x_j|$.

5. Randomized random walks

Let $X = (X_n : n \in \mathbb{N}_0)$ be a homogeneous irreducible Markov chain on a finite state space $\overline{F}$ with one-step transition probability matrix $r = (r(i,j) : i, j \in \overline{F})$ and (unique) steady-state distribution $\eta = (\eta_i : i \in \overline{F})$. This chain represents in our network applications the homogeneous Markov chain that describes the random walk (routing) of the customers on the nodes of the network. In this application scenario the routing will be modified by a network controller as a reaction to changes of the network's parameter due to the impact of the environment.

The general principle is: The transition matrix r will be used as a "candidate generating matrix" for the next state of the random walk. The candidate state will be accepted with state dependent probabilities and we develop different policies to continue when the proposed state is rejected.

5.1. Randomized skipping

The following modification of X with prescribed $B \subsetneq \overline{F}$ is known under the terms of "Markov chains with taboo set B", or "jump-over protocol for B", or "skipping B". In the realm of queueing network theory this principle for modifying a Markov chain seems to be introduced independently several times and was used to resolve blocking, see e.g. [Dij88], [EF98] and [Ser99, Chapter 3.6], (where it is called blocking and rerouting) and the references therein. As a general methodology skipping was already introduced by Schassberger [Sch84] and later on, based on Schassberger's result, it was used in [DS96] to construct general abstract network processes. For networks with unreliable nodes it was introduced in [EF98] and [SD03a].

An intuitive description of the skipping principle can be given in terms of a random walk on $\overline{F}$ governed by r: If the random walker's (RW) path governed by the Markov chain is restricted by a taboo set $B \subsetneq \overline{F}$, RW applies

Skipping B: If RW is in state $i \in \overline{F}$ and selects (with probability $r(i,j)$) its destination $j \in \overline{F} \setminus B$, this jump is allowed and immediately performed. If (with probability $r(i,k)$) he decides to jump to state $k \in B$, he only performs an imaginary jump to k, spending no time there, but jumping on immediately governed by the matrix r, i.e. with probability $r(k,l)$ he selects another possible successor state l; if $l \in \overline{F} \setminus B$, then the jump is performed immediately, but if $l \in B$, RW has to perform another random choice as if he would depart from l; and so on.

Our modification scheme for X and r is a randomized generalization of skipping. For the states in $\overline{F}$ are given "acceptance probabilities" by some vector $\boldsymbol{\alpha} = (\alpha_j \in [0,1] : j \in \overline{F})$. The random walker selects his itinerary under r and constraints $\boldsymbol{\alpha}$ by

Randomized skipping with acceptance probabilities $\boldsymbol{\alpha}$: If RW is in state $i \in \overline{F}$ and selects (with probability $r(i,j)$) its destination $j \in \overline{F}$, a Bernoulli experiment is performed with success (acceptance) probability α_j, independent of the past, given j. If the experiment is successful ($= 1$), this jump is accepted, immediately performed, and

RW settles down at j for at least one time slot. If the experiment is not successful ($=0$), this jump is not accepted and RW only performs an imaginary jump to j, spends no time there, and jumps on immediately, governed by the matrix r, i.e. with probability $r(j,l)$ he selects another possible successor state l; thereafter a Bernoulli experiment is performed with success (acceptance) probability α_l, independent of the past, given l. If the experiment is successful ($=1$), this jump is accepted, immediately performed, and RW settles down at l for at least one time slot. If the experiment is not successful ($=0$), this jump is not accepted, RW only performs an imaginary jump to l, spends no time there, and jumps on immediately according to r; and so on.

Example 5.1.1. If for $B \subsetneq \overline{F}$ we set $\alpha_j = 0$ if $j \in B$, and $\alpha_j = 1$ if $j \in \overline{F} \setminus B$, we have the skipping over taboo set B as described above: A jump to $j \in B$ is never accepted, while a proposed jump to $j \in \overline{F} \setminus B$, will be accepted with probability 1.

For general $\boldsymbol{\alpha}$ the set $B(\boldsymbol{\alpha}) = \{j \in \overline{F} : \alpha_j = 0\}$ is a taboo set for the "randomized skipping process".

5.1.1. Transition matrix

It is easy to see that this construction of a modified chain by randomized skipping generates a new homogeneous Markov chain, the transition matrix of which will be denoted by $r^{(\boldsymbol{\alpha})}$. For simplicity of presentation we will denote a Markov chain under this regime by $X^{(\boldsymbol{\alpha})}$.

To determine the transition probabilities $r^{(\boldsymbol{\alpha})}(i,j)$ we construct an auxiliary absorbing Markov chain $(X^{(A)}, Y^{(A)})$ with state space $\overline{F} \times \{0,1\}$ such that

- $X^{(A)}$ records the itinerary of RW during his (possibly many) imaginary jumps until its candidate state is accepted – if this happened RW settles down there forever, because the chain $(X^{(A)}, Y^{(A)})$ is absorbed,
- $Y^{(A)}$ indicates whether a candidate state is accepted or not,
- the states in $\overline{F} \times \{1\}$ are absorbing,
- initial states for $(X^{(A)}, Y^{(A)})$ are restricted to $\overline{F} \times \{0\}$, and therefore $Y^{(A)}$ stays at 0 until absorption of $(X^{(A)}, Y^{(A)})$ in $\overline{F} \times \{1\}$.

The transition probabilities for $(X^{(A)}, Y^{(A)})$ are for $i, j \in \overline{F}$ as long as at time n the candidate state is not accepted

$$P(X^{(A)}_{n+1} = j, Y^{(A)}_{n+1} = 1 | X^{(A)}_n = i, Y^{(A)}_n = 0) = r(i,j)\alpha_j = (r \cdot I_{\boldsymbol{\alpha}})_{ij}, \tag{5.1.1}$$

$$P(X^{(A)}_{n+1} = j, Y^{(A)}_{n+1} = 0 | X^{(A)}_n = i, Y^{(A)}_n = 0) = r(i,j)(1-\alpha_j) = (r \cdot I_{(1-\boldsymbol{\alpha})})_{ij}, \tag{5.1.2}$$

and thereafter
$$P(X^{(A)}_{n+1} = j, Y^{(A)}_{n+1} = 1 | X^{(A)}_n = i, Y^{(A)}_n = 1) = \delta_{ij}, \tag{5.1.3}$$

$$P(X^{(A)}_{n+1} = j, Y^{(A)}_{n+1} = 0 | X^{(A)}_n = i, Y^{(A)}_n = 1) = 0. \tag{5.1.4}$$

We denote the (P-a.s. finite) first entrance time into $\overline{F} \times \{1\}$ of $(X^{(A)}, Y^{(A)})$ by

$$\tau^{(A)} := \inf\{n \geq 1 | (X^{(A)}_n, Y^{(A)}_n) \in F \times \{1\}\} = \inf\{n \geq 1 | Y^{(A)}_n = 1\}. \tag{5.1.5}$$

Theorem 5.1.2. *For the Markov chain $X = (X_n : n \in \mathbb{N}_0)$ with transition matrix $r = (r(i,j) : i,j \in \overline{F})$ and a non zero vector $\boldsymbol{\alpha} = (\alpha_j : j \in \overline{F})$ of acceptance probabilities denote by $X^{(\boldsymbol{\alpha})}$ the Markov chain modification of X under randomized skipping and by $r^{(\boldsymbol{\alpha})}$ the transition matrix of $X^{(\boldsymbol{\alpha})}$, and by $B(\boldsymbol{\alpha}) = \{j \in \overline{F} : \alpha_j = 0\} \subsetneq \overline{F}$ the taboo set for $X^{(\boldsymbol{\alpha})}$. Then from the auxiliary chain $(X^{(A)}, Y^{(A)})$ we obtain*

$$r^{(\boldsymbol{\alpha})}(i,j) = P\left(X^{(A)}(\tau^{(A)}) = j | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\right), \quad i,j \in \overline{F}. \tag{5.1.6}$$

The Markov chain $X^{(\boldsymbol{\alpha})}$ with state space $\overline{F}$ and transition matrix $r^{(\boldsymbol{\alpha})}$ is irreducible on $\overline{F} \setminus B(\boldsymbol{\alpha})$, and it holds

$$r^{(\boldsymbol{\alpha})} = \sum_{k=0}^{\infty} \left(rI_{(1-\boldsymbol{\alpha})}\right)^k rI_{\boldsymbol{\alpha}} = \left(I - rI_{(1-\boldsymbol{\alpha})}\right)^{-1} rI_{\boldsymbol{\alpha}}. \tag{5.1.7}$$

Proof. Irreducibility of $X^{(\boldsymbol{\alpha})}$ on $\overline{F} \setminus B(\boldsymbol{\alpha})$ and the first statement (5.1.6) follows directly from the construction.

For $i \in \overline{F}, j \in B(\boldsymbol{\alpha})$ it follows $r^{(\boldsymbol{\alpha})}(i,j) = 0$ from the definition of acceptance probability. Note, that from (5.1.2) follows

$$\begin{aligned} P\Big(X_n^{(A)} = k, Y_n^{(A)} = 0, Y_{n-1}^{(A)} = 0, \dots \\ \dots, Y_2^{(A)} = 0, Y_1^{(A)} = 0, | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\Big) = \left(\left(rI_{(1-\boldsymbol{\alpha})}\right)^n\right)_{ik}. \end{aligned} \tag{5.1.8}$$

It follows for $i \in \overline{F}, j \in \overline{F} \setminus B(\boldsymbol{\alpha})$

$$\begin{aligned} r_{ij}^{(\boldsymbol{\alpha})} &= P\left(X^{(A)}(\tau^{(A)}) = j | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\right) \\ &= P\left(X^{(A)}(\tau^{(A)}) = j, Y^{(A)}(\tau^{(A)}) = 1 | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\right) \\ &= \sum_{n=1}^{\infty} P\left(X^{(A)}(\tau^{(A)}) = j, Y^{(A)}(\tau^{(A)}) = 1, \tau^{(A)} = n | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\right) \\ &= \sum_{n=1}^{\infty} \sum_{k \in \overline{F}} P\Big(X_n^{(A)} = j, Y_n^{(A)} = 1, X_{n-1}^{(A)} = k, Y_{n-1}^{(A)} = 0, Y_{n-2}^{(A)} = 0, \dots \\ &\qquad\qquad \dots, Y_2^{(A)} = 0, Y_1^{(A)} = 0 | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\Big) \\ &\overset{(*)}{=} \sum_{n=1}^{\infty} \sum_{k \in \overline{F}} P\left(X_n^{(A)} = j, Y_n^{(A)} = 1 | X_{n-1}^{(A)} = k, Y_{n-1}^{(A)} = 0\right) \\ &\qquad \cdot P\Big(X_{n-1}^{(A)} = k, Y_{n-1}^{(A)} = 0, Y_{n-2}^{(A)} = 0, \dots \\ &\qquad\qquad \dots, Y_2^{(A)} = 0, Y_1^{(A)} = 0 | (X_0^{(A)}, Y_0^{(A)}) = (i,0)\Big) \end{aligned}$$

$$\begin{aligned} &\overset{(**)}{=} \sum_{k \in \overline{F}} \sum_{n=1}^{\infty} (rI_{\boldsymbol{\alpha}})_{kj} \left(\left(rI_{(1-\boldsymbol{\alpha})}\right)^{n-1}\right)_{ik} \overset{(***)}{=} \sum_{k \in \overline{F}} (rI_{\boldsymbol{\alpha}})_{kj} \left(\sum_{n=1}^{\infty} \left(rI_{(1-\boldsymbol{\alpha})}\right)^{n-1}\right)_{ik} \\ &= \left(\left(I - rI_{(1-\boldsymbol{\alpha})}\right)^{-1} rI_{\boldsymbol{\alpha}}\right)_{ij}, \end{aligned}$$

which is (5.1.7). Here (*) utilizes the Markov property of $(X^{(A)}, Y^{(A)})$, (**) follows from (5.1.8) and (5.1.1), and in (***) convergence of $\sum_{n=1}^{\infty} \left(rI_{(1-\boldsymbol{\alpha})}\right)^{n-1}$ follows from irreducibility of r and from the substochasticity of $rI_{(1-\boldsymbol{\alpha})}$, which is strict for at least one row. □

We note that the states in $B(\boldsymbol{\alpha})$ are inessential, because the $B(\boldsymbol{\alpha})$-columns $(r_{ij}^{(\boldsymbol{\alpha})} : i \in \overline{F})$ for $j \in B(\boldsymbol{\alpha})$ are zero.

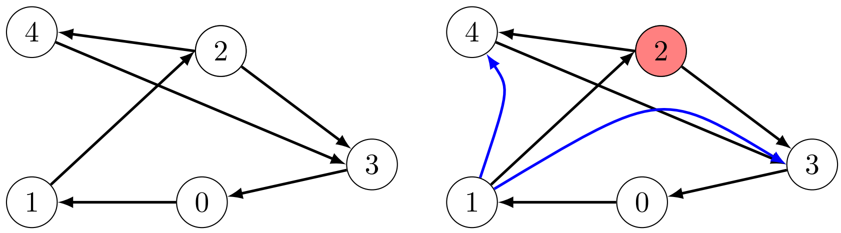
Example 5.1.3. Consider a set $\overline{F} = \{0, 1, 2, 3, 4\}$ and a routing matrix

$$r = \left(\begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & & 1 & & & \\ 1 & & & 1 & & \\ 2 & & & & 0.6 & 0.4 \\ 3 & 1 & & & & \\ 4 & & & & 1 & \end{array}\right).$$

After applying the skipping rule to r with availability vector $\boldsymbol{\alpha} = (1, 1, 0.5, 1, 1)$ we get

$$r^{(\boldsymbol{\alpha})} = \left(\begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & & 1 & & & \\ 1 & & & 0.5 & 0.3 & 0.2 \\ 2 & & & & 0.6 & 0.4 \\ 3 & 1 & & & & \\ 4 & & & & 1 & \end{array}\right).$$

See Figure 5.1.1 on page 154.



(a) original matrix r (b) modified matrix $r^{(\boldsymbol{\alpha})}$

Figure 5.1.1.: Matrix r on $\overline{F} = \{0, 1, 2, 3, 4\}$ and $r^{(\boldsymbol{\alpha})}$ with $\boldsymbol{\alpha} = (1, 1, 0.5, 1, 1)$ from Example 5.1.3.

5.1.2. Stationary distribution

Recall that X is irreducible with transition matrix r on the finite state space $\overline{F}$ and has the unique stationary distribution $\eta = (\eta_j : j \in \overline{F})$. For a non zero vector $\boldsymbol{\alpha} = (\alpha_j : j \in \overline{F})$ the chain $X^{(\boldsymbol{\alpha})}$ with state space $\overline{F}$ and transition matrix $r^{(\boldsymbol{\alpha})}$ is irreducible on $\overline{F} \setminus B(\boldsymbol{\alpha})$, with $B(\boldsymbol{\alpha}) = \{j \in \overline{F} : \alpha_j = 0\}$. We denote its stationary distribution by $\eta^{(\boldsymbol{\alpha})} = (\eta^{(\boldsymbol{\alpha})}(j) : j \in \overline{F})$, which has support $\overline{F} \setminus B(\boldsymbol{\alpha})$. We will study the relation between η and $\eta^{(\boldsymbol{\alpha})}$.

Proposition 5.1.4. *Let x be a solution of the balance equation $x \cdot r = x$ then $y = x \cdot I_{\boldsymbol{\alpha}}$ solves the steady-state equation $y \cdot r^{(\boldsymbol{\alpha})} = y$ of the modified Markov chain with randomized skipping.*

Proof. From $x \cdot r = x$ we obtain

$$xr\underbrace{\left(I_{\boldsymbol{\alpha}} + I_{(1-\boldsymbol{\alpha})}\right)}_{=I} = x \Longleftrightarrow xrI_{\boldsymbol{\alpha}} = x - xrI_{1-\boldsymbol{\alpha}} \Longleftrightarrow xrI_{\boldsymbol{\alpha}} = x(I - rI_{(1-\boldsymbol{\alpha})})$$

$$\Longleftrightarrow x\underbrace{(I - rI_{(1-\boldsymbol{\alpha})})(I - rI_{(1-\boldsymbol{\alpha})})^{-1}}_{=I}rI_{\boldsymbol{\alpha}} = x(I - rI_{(1-\boldsymbol{\alpha})})$$

and with $y := x(I - rI_{(1-\alpha)})$ follows that a required solution of $y \cdot r^{(\boldsymbol{\alpha})} = y$ is

$$y \quad = \quad x(I - rI_{(1-\boldsymbol{\alpha})}) = x - xI_{(1-\boldsymbol{\alpha})} = xI_{\boldsymbol{\alpha}}.$$

□

Proposition 5.1.5. *Let y be a solution of the balance equation $y \cdot r^{(\boldsymbol{\alpha})} = y$ then*

$$x = y(I - rI_{(1-\boldsymbol{\alpha})})^{-1} \tag{5.1.9}$$

is a solution of the steady-state equation $x \cdot r = x$ and it holds

$$y = (\alpha_j x_j : j \in \overline{F}). \tag{5.1.10}$$

Proof. From the definition of $r^{(\boldsymbol{\alpha})}$ follows

$$y\underbrace{(I - rI_{(1-\boldsymbol{\alpha})})^{-1}rI_{\boldsymbol{\alpha}}}_{=r^{(\boldsymbol{\alpha})}} = y\underbrace{(I - rI_{(1-\boldsymbol{\alpha})})^{-1}(I - rI_{(1-\boldsymbol{\alpha})})}_{=I}.$$

So $x = y(I - rI_{(1-\boldsymbol{\alpha})})^{-1}$ fulfills

$$xrI_{\boldsymbol{\alpha}} = x(I - rI_{(1-\boldsymbol{\alpha})}) \Longleftrightarrow xr\underbrace{\left(I_{\boldsymbol{\alpha}} + I_{(1-\boldsymbol{\alpha})}\right)}_{=I} = x$$

which is $x \cdot r = x$. (5.1.10) follows from $x(I - rI_{(1-\boldsymbol{\alpha})}) = y$ as in Proposition 5.1.4. □

Theorem 5.1.6. *If η is the unique steady-state distribution of the irreducible Markov chain X on finite state space $\overline{F}$, then the unique steady-state distribution $\eta^{(\boldsymbol{\alpha})}$ of the Markov chain $X^{(\boldsymbol{\alpha})}$ is, with normalization constant $C^{(\boldsymbol{\alpha})} = (\eta I_{\boldsymbol{\alpha}}\mathbf{e}) = \langle \eta, \boldsymbol{\alpha} \rangle$ and support $\overline{F} \setminus B(\boldsymbol{\alpha})$,*

$$\eta^{(\boldsymbol{\alpha})} = \left(C^{(\boldsymbol{\alpha})}\right)^{-1}(\eta_j \alpha_j : j \in \overline{F}). \tag{5.1.11}$$

Proof. From Proposition 5.1.4 and Proposition 5.1.5 we know that there is a one-to-one connection between all solutions of $x \cdot r = x$ and $y \cdot r^{(\boldsymbol{\alpha})} = y$.

From the assumptions on X we know that η is the unique stochastic solution of $x \cdot r = x$. Its normalized companion is therefore (5.1.11) by (5.1.10). □

5.2. Randomized reflection

An important problem in the control of transmission networks is to react to full buffers at receiver stations by the network provider. There are many detailed control regimes to resolve blocking which occurs when buffers overflow. It turned out that it is usually difficult to construct analytical network models for this problem which admit closed form solutions for main performance metrics.

The following control principle is common to resolve blocking situations. In networks with blocking of stations due to full buffers or blocking due to resource sharing it is called blocking principle REPETITIVE SERVICE – RANDOM DESTINATION (RS–RD).

For applications in modeling of communication protocols in systems with finite buffers or for ALOHA-type protocols see [Kle76, Section 5.11]. Within the abstract framework of reversible processes it occurs in [Lig85, Proposition II.5.10]. For networks with unreliable nodes it was introduced in [EF98] and [SD03a].

The principle is in case of full buffer regulation, that whenever a packet is sent from some node i to node j and it is observed that the buffer for incoming packets at j is full, the packet is rejected (lost) and node i, who has saved a copy, tries to resend this packet (REPETITIVE SERVICE), but not necessarily to j (RANDOM DESTINATION). This procedure is iterated until the packet is sent to a node with free buffer places.

We will apply this scheme to modify the homogeneous irreducible Markov chain $X = (X_n : n \in \mathbb{N}_0)$ on the finite state space $\overline{F}$ with one-step transition probability matrix $r = (r(i,j) : i, j \in \overline{F})$ and (unique) steady-state distribution $\eta = (\eta_i : i \in \overline{F})$.

In terms of a random walk on $\overline{F}$ governed by r an intuitive description is: If the random walker's (RW) path governed by the Markov chain is restricted by a taboo set $B \subsetneq \overline{F}$, RW applies

Reflection at B: If RW is in state $i \in \overline{F}$ and selects (with probability $r(i,j)$) its destination $j \in \overline{F} \setminus B$, this jump is allowed and immediately performed. If (with probability $r(i,k)$) he decides to jump to state $k \in B$, he is reflected at k and spends at least one further time slot at node i. Thereafter he restarts the procedure possibly with another destination node, i.e. with probability $r(i,l)$ he selects successor state l; if $l \in \overline{F} \setminus B$ then the jump is performed immediately, but if $l \in B$ RW is reflected again; and so on.

Our extended modification scheme for X and r is a randomized generalization of that reflection scheme. For the states in $\overline{F}$ are given "acceptance probabilities" by some vector $\boldsymbol{\alpha} = (\alpha_j \in [0,1] : j \in \overline{F})$. The random walker (RW) selects his itinerary under r and the constraints $\boldsymbol{\alpha}$ by

Randomized reflection with acceptance probabilities $\boldsymbol{\alpha}$: If RW is in state $i \in \overline{F}$ and selects (with probability $r(i,j)$) its destination $j \in \overline{F}$, a Bernoulli experiment is performed with success (acceptance) probability α_j, independent of the past, given j. If the experiment is successful (= 1), this jump is accepted, immediately performed, and RW settles down at j for at least one time slot. If the experiment is not successful (= 0), this jump is not accepted and RW stays on at i for at least one further time slot. If this slot expires with probability $r(i,l)$ RW selects another possible successor state l; thereafter a Bernoulli experiment is performed with success (acceptance) probability α_l, independent of the past, given l. If the experiment is successful (= 1), this jump is accepted, immediately performed, and RW settles down at l for at least one time slot. If the experiment is not successful (= 0), this jump is not accepted and RW stays on at i for at least one further time slot; and so on.

Example 5.2.1. If for $B \subsetneq \overline{F}$ we set $\alpha_j = 0$ if $j \in B$, and $\alpha_j = 1$ if $j \in \overline{F} \setminus B$, we have the reflection at taboo set B as described above: A jump to $j \in B$ is never accepted, while a proposed jump to $j \in \overline{F} \setminus B$, will be accepted with probability 1.

For general $\boldsymbol{\alpha}$ the set $B(\boldsymbol{\alpha}) = \{j \in \overline{F} : \alpha_j = 0\}$ is a taboo set for the "randomized reflection process".

5.2.1. Transition matrix and stationary distribution

It is easy to see that the construction of a process by randomized reflection generates a homogeneous Markov chain.
For the Markov chain $X = (X_n : n \in \mathbb{N}_0)$ with transition matrix $r = (r(i,j) : i,j \in \overline{F})$ and a non zero vector $\boldsymbol{\alpha} = (\alpha_j : j \in \overline{F})$ of acceptance probabilities denote by $X^{(\boldsymbol{\alpha})}$ the Markov chain modification of X under randomized reflection and by $r^{(\boldsymbol{\alpha})}$ the transition matrix of $X^{(\boldsymbol{\alpha})}$, and by $B(\boldsymbol{\alpha}) = \{j \in \overline{F} : \alpha_j = 0\} \subsetneq \overline{F}$ the taboo set for $X^{(\boldsymbol{\alpha})}$. Then

$$r^{(\boldsymbol{\alpha})}(i,j) = \begin{cases} r(i,j)\alpha_j, & i,j \in \overline{F},\ i \neq j, \\ r(i,i) + \sum_{k \in \overline{F}} r(i,k)(1-\alpha_j), & i \in \overline{F},\ i = j. \end{cases} \tag{5.2.1}$$

$X^{(\boldsymbol{\alpha})}$ with state space $\overline{F}$ and transition matrix $r^{(\boldsymbol{\alpha})}$ may be reducible even on $\overline{F} \setminus B(\boldsymbol{\alpha})$.

A standard assumption in the literature for applying this reflection principle as rerouting scheme is that the original routing Markov chain X, resp. its transition matrix r is reversible for some probability measure $\eta = (\eta_i : i \in \overline{F})$. We shall set this assumption always in force when investigating this protocol. Determining the modified transition matrix $r^{(\boldsymbol{\alpha})}$ and its stationary distribution $\eta^{(\boldsymbol{\alpha})}$ is direct in this case.

Proposition 5.2.2. *If η is the unique steady-state distribution of the reversible irreducible Markov chain X on the finite state space $\overline{F}$, then the Markov chain $X^{(\boldsymbol{\alpha})}$ is reversible with steady-state distribution $\eta^{(\boldsymbol{\alpha})}$ with normalization constant $C^{(\boldsymbol{\alpha})} = (\eta I_{\boldsymbol{\alpha}} \mathbf{e}) = \langle \eta, \boldsymbol{\alpha} \rangle$ and support $\overline{F} \setminus B(\boldsymbol{\alpha})$:*

$$\eta^{(\boldsymbol{\alpha})} = \left(C^{(\boldsymbol{\alpha})}\right)^{-1} (\eta_j \alpha_j : j \in \overline{F}). \tag{5.2.2}$$

Proof. The proof is by directly checking the detailed balance equations for $r^{(\boldsymbol{\alpha})}$:

$$(\eta_j \alpha_j) r^{(\boldsymbol{\alpha})}(j,i) = (\eta_i \alpha_i) r^{(\boldsymbol{\alpha})}(i,j) \Leftrightarrow (\eta_j \alpha_j) r(j,i) \alpha_i = (\eta_i \alpha_i) r(i,j) \alpha_j, \qquad i,j \in \overline{F}.$$

□

5.3. Discussion of randomization algorithms

A closer look on our randomization procedures, considered as algorithms to manipulate a random walk and its stationary distribution, reveals close connections to standard simulation procedures. This will be discussed in this section and we will compare furthermore the randomization procedures.

5.3.1. General sampling schemes

Our starting point is an ergodic random walk X on a finite set $\overline{F}$ with stationary distribution η. X is modified according to an acceptance regime $\boldsymbol{\alpha}$. Our aim is to study the impact of this modification on the stationary distribution. This is different to the standard problem of simulation, i.e. to generate samples from a given distribution, which is known in principle, but not easy to access. We mention two simulation algorithms which are structurally related to our procedures.

(I) Markov Chain Monte Carlo algorithms start from a target distribution η which is usually not directly accessible for which e.g. $\int_{\overline{F}} f(x)\eta(dx) = \sum_{x\in\overline{F}} f(x)\eta_x$ has to be computed. The idea is to construct a homogeneous Markov chain with stationary distribution η and to sample from the Markov chains' state distribution after a sufficiently long time horizon. The construction of the one-step transition kernel following Hastings or Metropolis (see [Bre99, Section 7.7.1]) yields a two-step scheme:

1. From the present state generate a proposal for the next state and decide about acceptance of the proposed state,
2. if the proposal is rejected stay on at the present state and restart after a time slot.

This is a procedure as described in Proposition 5.2.2. The conclusion is that the rerouting scheme **rs-rd** used to model transmission networks' reaction to full buffers and its generalization, our randomized reflection, can be seen as MCMC processes.

(II) Von Neumann's acceptance-rejection method for sampling from a complicated distribution η (see [Bre99, p. 292]) is related to randomized skipping. If the random walk of Section 5.1 is an i.i.d. sequence the randomized skipping procedure is exactly sampling from $(\eta_j\alpha_j : j \in \overline{F})$ by sampling from η with possible rejection.

We remark that in MCMC algorithms usually reversible chains are constructed with acceptance probabilities which may depend on the departure state and the proposal. Such generalization is easily constructed here as well, in our notation we would incorporate success probabilities α_{ij}. Proposition 5.2.2 can be modified directly. A similar property and its proof for randomized skipping for non-reversible Markov chains seems to be not possible without loosing the simple to evaluate steady-state distribution.

5.3.2. Importance sampling

Our randomization procedures resemble obviously importance sampling procedures in simulations, because we

- produce a weighted version of the probability distribution η on $\overline{F}$, and
- define a modified random walk that generates this weighted distribution as its limiting distribution,

and this can therefore be considered as construction of a change of measure. Such change of measure is used to make the "more important states" more often visited by the random walker. An important problem is to compute expectations of the form

$$\int_{\overline{F}} f(x)\eta(dx) = \sum_{x\in\overline{F}} f(x)\eta_x, \tag{5.3.1}$$

where $X = (X_n : n \in \mathbb{N}_0)$ with transition matrix r is ergodic with limiting distribution η, where η is not known or not easily accessible.

For simplicity we assume $\alpha_j > 0\ \forall j \in \overline{F}$ and use the notation of the previous sections. Recall the normalization constant $C^{(\boldsymbol{\alpha})} = \langle \eta, \boldsymbol{\alpha} \rangle$ of $\eta^{(\boldsymbol{\alpha})}$. By simple manipulation

$$\sum_{x\in\overline{F}} f(x)\eta_x = \sum_{x\in\overline{F}} \frac{f(x)}{\langle \eta, \boldsymbol{\alpha}\rangle^{-1}} \cdot \frac{\eta_x}{\langle \eta, \boldsymbol{\alpha}\rangle} = \frac{\sum_{x\in\overline{F}} \frac{f(x)}{\alpha_x} \cdot \frac{\eta_x\alpha_x}{\langle \eta, \boldsymbol{\alpha}\rangle}}{\sum_{x\in\overline{F}} \frac{1}{\alpha_x} \cdot \frac{\eta_x\alpha_x}{\langle \eta, \boldsymbol{\alpha}\rangle}}.$$

Nominator and denominator are integrals with respect to the stationary distribution of $X^{(\boldsymbol{\alpha})}$ derived from X with acceptance probability $\boldsymbol{\alpha} = (\alpha_x : x \in \overline{F})$. X is ergodic and $\alpha_x > 0\ \forall x \in \overline{F}$ implies that $X^{(\boldsymbol{\alpha})}$ is ergodic on $\overline{F}$ as well.

Denote by $T(i)$ the first entrance time into i, then from the regenerative structure of $X^{(\boldsymbol{\alpha})}$ we have

$$\sum_{x\in\overline{F}} \frac{f(x)}{\alpha_x} \cdot \frac{\eta_x\alpha_x}{\langle \eta, \boldsymbol{\alpha}\rangle} = \frac{E^{(\boldsymbol{\alpha})}\left[\sum_{n=0}^{T(i)-1} \frac{f(X_n^{(\boldsymbol{\alpha})})}{\alpha_{X_n^{(\boldsymbol{\alpha})}}} \middle| X_0^{(\boldsymbol{\alpha})} = i\right]}{E^{(\boldsymbol{\alpha})}\left[T(i) \middle| X_0^{(\boldsymbol{\alpha})} = i\right]}$$

and

$$\sum_{x\in\overline{F}} \frac{1}{\alpha_x} \cdot \frac{\eta_x\alpha_x}{\langle \eta, \boldsymbol{\alpha}\rangle} = \frac{E^{(\boldsymbol{\alpha})}\left[\sum_{n=0}^{T(i)-1} \frac{1}{\alpha_{X_n^{(\boldsymbol{\alpha})}}} \middle| X_0^{(\boldsymbol{\alpha})} = i\right]}{E^{(\boldsymbol{\alpha})}\left[T(i) \middle| X_0^{(\boldsymbol{\alpha})} = i\right]}.$$

This yields eventually

$$\sum_{x\in\overline{F}} f(x)\eta_x = \frac{E^{(\boldsymbol{\alpha})}\left[\sum_{n=0}^{T(i)-1} \frac{f(X_n^{(\boldsymbol{\alpha})})}{\alpha_{X_n^{(\boldsymbol{\alpha})}}} \middle| X_0^{(\boldsymbol{\alpha})} = i\right]}{E^{(\boldsymbol{\alpha})}\left[\sum_{n=0}^{T(i)-1} \frac{1}{\alpha_{X_n^{(\boldsymbol{\alpha})}}} \middle| X_0^{(\boldsymbol{\alpha})} = i\right]}.$$

To obtain an estimator for $E^{(\boldsymbol{\alpha})}\left[\sum_{n=0}^{T(i)-1} g\left(X_n^{(\boldsymbol{\alpha})}\right) \middle| X_0^{(\boldsymbol{\alpha})} = i\right]$ denote $T^0(i) \equiv 0$, and $T^{(n+1)}(i) := \inf(k > T^{(n)}(i) : X_k^{(\boldsymbol{\alpha})} = i)$, $n \geq 0$, and take the sample means of the independent replications

$$\frac{1}{K}\sum_{k=1}^{K}\left(\sum_{n=T^{(k-1)}(i)}^{T^{(k)}(i)} g\left(X_n^{(\boldsymbol{\alpha})}\right) \middle| X_0^{(\boldsymbol{\alpha})} = i\right),$$

which converge a.s. to the target expression for $K \to \infty$.

So, if it is possible to simulate without too much effort the system described by the Markov chain $X^{(\boldsymbol{\alpha})}$, we can approximate the integral by the quotient of the two time averages over the regeneration periods.

Moreover, additionally to making important states more probable to visit, for a given initial state i we can control the expected length of the regeneration period.

If X is reversible, the one step transition matrix of the chain modified by randomized reflection is directly accessible, see (5.2.1).

In the general case when X is modified by randomized skipping the one step transition matrix of the modified chain is possibly not directly available, see (5.1.7). We can remedy this drawback by simulating instead the auxiliary chain $(X^{(A)}, Y^{(A)})$ (with obvious modifications) which is used to find the one step transition matrix for randomized skipping in Section 5.1.1. The modification of $(X^{(A)}, Y^{(A)})$ is not to make $\overline{F} \times \{1\}$ absorbing. The transition rates (5.1.3) and (5.1.4) are replaced by

$$P(X_{n+1}^{(A)} = j, Y_{n+1}^{(A)} = 1 | X_n^{(A)} = i, Y_n^{(A)} = 1) = r(i,j)\alpha_j, \tag{5.3.2}$$

$$P(X_{n+1}^{(A)} = j, Y_{n+1}^{(A)} = 0 | X_n^{(A)} = i, Y_n^{(A)} = 1) = r(i,j)(1-\alpha_j). \tag{5.3.3}$$

In the time average over a regeneration period we then have to sum instead of $f(X_n^{(\boldsymbol{\alpha})}(\omega))$ the values $f(X_n^{(A)}(\omega)) \cdot Y_n^{(A)}(\omega)$ and instead of $\frac{1}{\alpha_{X_n^{(\boldsymbol{\alpha})}(\omega)}}$ the values $\frac{Y_n^{(A)}(\omega)}{\alpha_{X_n^{(\boldsymbol{\alpha})}(\omega)}}$ and to divide by the expected sojourn time in $\overline{F} \times \{1\}$ until return to $(i,1)$ given start in $(i,1)$ (which cancels out).

5.3.3. Comparison of randomized skipping and randomized reflection

We can compare transition matrices having the same dimension and stationary distribution by Peskun ordering (see [Pes73]) which is standard ordering in MCMC simulations.

Definition 5.3.1. Let $r = (r(i,j) : i,j \in \overline{F})$ and $r' = (r'(i,j) : i,j \in \overline{F})$ be transition matrices on a finite set $\overline{F}$ such that $\xi r = \xi r' = \xi$ for a probability vector ξ.

We say that r' is smaller than r in the Peskun order, $r' \prec_P r$, if for all $j, i \in \overline{F}$ with $i \neq j$ holds $r'(j,i) \leq r(j,i)$.

Peskun used this order to compare reversible transition matrices with the same stationary distribution and their asymptotic variance. A useful interpretation is that in case of $r' \prec_P r$ a random walker under r explores his state space faster than under r'.

Proposition 5.3.2. *Let X be irreducible with transition matrix r on the finite state space $\overline{F}$ which is reversible for the stationary distribution $\eta = (\eta_j : j \in \overline{F})$, and $\boldsymbol{\alpha} = (\alpha_j : j \in \overline{F})$ be a non zero vector.*

Denote by $r_{skip}^{(\boldsymbol{\alpha})}$ the modification of r by randomized skipping and by $r_{refl}^{(\boldsymbol{\alpha})}$ the modification of r by randomized reflection. Then it holds $r_{refl}^{(\boldsymbol{\alpha})} \prec_P r_{skip}^{(\boldsymbol{\alpha})}$.

Claim. The proof is obvious because for all $j, i \in \overline{F}$ with $i \neq j$ holds $r_{refl}^{(\boldsymbol{\alpha})}(i,j) = r(i,j) \cdot \alpha_j$ and $r_{skip}^{(\boldsymbol{\alpha})}(i,j) = r(i,j) \cdot \alpha_j +$ possible further terms, which are all non-negative.

Consequences of $r_{refl}^{(\boldsymbol{\alpha})} \prec_P r_{skip}^{(\boldsymbol{\alpha})}$ in Proposition 5.3.2 are

1. the asymptotic variance in the central limit theorem for a Markov chain driven by $r_{skip}^{(\boldsymbol{\alpha})}$ is smaller than for a Markov chain driven by $r_{refl}^{(\boldsymbol{\alpha})}$ although both chains have the same limiting distribution (see [Tie98]).

2. the spectral gap of a Markov chain driven by $r_{refl}^{(\boldsymbol{\alpha})}$ is smaller than for a Markov chain driven by $r_{skip}^{(\boldsymbol{\alpha})}$ although both chains have the same limiting distribution. This means that the modification of $\mathbf{X}$ by randomized skipping converges faster to equilibrium than the modification of $\mathbf{X}$ by randomized reflection.

A further distinction between modifications of $\mathbf{X}$ by randomized skipping and reflection is visible if we consider for a reversible Markov chain $\mathbf{X}$ with transition matrix r the associated transition graph $G = (\overline{F}, \mathcal{E})$ where $(i,j) \in \mathcal{E} \Leftrightarrow r(i,j) > 0$.

Note that from reversibility follows $(i,j) \in \mathcal{E} \Leftrightarrow (j,i) \in \mathcal{E}$. Denote by $\mathcal{N}(i) := \{j \in \overline{F} : (i,j) \in \mathcal{E}\}$ the one-step neighborhood of i.

If we consider randomized skipping and randomized reflection as algorithms to produce modifications of the random walk $\mathbf{X}$, then the algorithm for randomized reflection is local with respect to G because transitions out of states i are determined only on the basis of decisions in $\mathcal{N}(i)$.

On the other side, randomized skipping is global with respect to G because for transitions out of i we possibly must use random experiments anywhere on the graph.

Construction of local algorithms for control, scheduling, and development of complex systems is important for key applications, e.g. wireless sensor networks or autonomous interacting robotic systems. For mathematical investigations of this problem we refer to [SS12], where an optimal local control algorithm for wireless sensor networks is determined by local approximation of an optimal global algorithm.

6. Modified Jackson networks

We consider a Jackson network [Jac57] with node set $\overline{J} := \{1, \dots, J\}$. Customers arrive in independent external Poisson streams, at node j with finite intensity $\lambda_j \geq 0$, we set $\lambda = \lambda_1 + \ldots + \lambda_J > 0$. Customers are indistinguishable and follow the same rules. Customers request for service which is exponentially distributed with mean 1 at all nodes, all these requests constitute an independent family of variables which are independent of the arrival streams.

Nodes are exponential single servers with state dependent service rates and infinite waiting room under first-come-first-served (FCFS) regime. If at node j are $n_j > 0$ customers, either in service or waiting, service is provided there with intensity $\mu_j(n_j) > 0$. Routing is Markovian, a customer departing from node i immediately proceeds to node j with probability $r(i,j) \geq 0$, and departs from the network with probability $r(j,0)$. Taking $r(0,j) = \lambda_j/\lambda,\ r(0,0) = 0$, we assume that the extended routing matrix $r = (r(i,j) : i,j = 0,\dots,J)$ is irreducible. Then the traffic equations

$$\eta_j = \lambda_j + \sum_{i=1}^{J} \eta_i r(i,j), \qquad j = 1,\dots,J, \tag{6.0.1}$$

have a unique solution which we denote by $\eta = (\eta_j : j = 1,\dots,J)$. We extend the traffic equation (6.0.1) to a steady-state equation for a routing Markov chain by

$$\eta_j = \sum_{i=0}^{J} \eta_i r(i,j), \qquad j = 0,1,\dots,J, \tag{6.0.2}$$

which is solved by $\eta = (\eta_j : j = 0,1,\dots,J)$, where $\eta_0 := \lambda$, the other η_j are from (6.0.1).

We use η in both meanings and emphasize the latter one by *extended traffic solution* η. η is usually not a stochastic vector. Let $\mathbf{X} = (X(t) : t \geq 0)$ denote the vector process recording the queue lengths in the network for time t. $X(t) = (X_1(t),\dots,X_J(t)) \in \mathbb{N}_0^{\overline{J}}$ reads: at time t there are $X_j(t)$ customers present at node j, either in service or waiting. The assumptions put on the system imply that $\mathbf{X}$ is a strong Markov process on state space $\mathbb{N}_0^{\overline{J}}$, we denote its generator $Q^{\mathbf{X}} = (q^{\mathbf{X}}(\mathbf{n},\mathbf{n}') : \mathbf{n},\mathbf{n}' \in \mathbb{N}_0^{\overline{J}})$.

For an ergodic network process $\mathbf{X}$ Jackson's theorem [Jac57] states that the unique steady-state and limiting distribution ξ on $\mathbb{N}_0^{\overline{J}}$ is with normalizing constants $C(j)$ for the marginal (over nodes) distributions

$$\xi(\mathbf{n}) = \xi(n_1,\dots,n_J) = \prod_{j=1}^{J}\prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} C(j)^{-1}. \tag{6.0.3}$$

6.1. Modifications: Upgraded and/or degraded service and adapted routing

Optimal design of queueing systems is a challenging operations and research field, a recent survey is the book of Stidham [Sti09]. One of the typical problem structures is as follows: Given the nodes of the network and their service capacities and the external arrival streams (load) to the network. Find an "optimal routing" for the customers such that some optimality criterion is satisfied.
In [BS83] a set of stations with fixed service intensities μ_i is located in parallel, and a Poisson arrival stream with overall load λ has to be split and routed to the stations such that a given cost-reward function is maximized. This results in determining the routing of arriving customers and the "optimal utilizations" $\lambda \cdot r(0,i)/\mu_i$ (with self-explaining notation). In [Sti09, Sect. 6.1] in addition the overall load rate λ is subject to control, and in [Sti09, Chapt. 7] the design of general Jackson networks by control of admission and routing is investigated.

With a slightly different setting of the Jackson networks, resp. a more general migration network, described by parameters which are essentially equivalent to ours, Whittle [Whi85, Sect. 2] considered optimal "non-adaptive routing rules" (not depending on the queue lengths as in our model as well). Optimization of routing for fixed external arrival intensities then results (in our notation) in optimal routing probabilities $r(i,j)$ and in the cost function occur the optimal ratios $\eta_i/\mu_i(n_i)$, resp, $\eta_i^{n_i}/(\prod_{\ell=1}^{n_i} \mu_i(\ell))$.

In this section we assume that according to some optimization criterion with respect to the local queue lengths X_j for the fixed $\mu_j(n_j)$, $j \in \overline{J}$, of the Jackson network the optimal utilizations η_j/μ_j, resp, the ratios $\eta_j/\mu_j(n_j)$ are determined for fixed service rates $\mu_j(n_j)$ by an adequate routing scheme r.
We investigate the problem how to adjust routing when service capacities change, such that the ratios $\eta_j/\mu_j(n_j)$ are maintained. This would guarantee that the optimal local queue lengths are maintained optimally according to the original criterion.
If the service intensities $\mu_i(n_i)$ at node i are changed by a factor $\gamma_i \in [0,\infty)$ for $i \in \overline{J}$, we have to react in different ways depending on the size of the γ_i.
It is possible that some nodes can break down completely, i.e. $\gamma_\ell = 0$ for such node ℓ. Clearly, the broken down nodes should not be visited any longer, and from the side constraint to maintain at least approximately the ratios "overall arrival rate"/"service rates", it is tempting to try rerouting by randomized skipping or reflection to some suitably selected "acceptance probability vector" according to some suitable $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\boldsymbol{\gamma})$. (If there is no ambiguity we will shortly write only $\boldsymbol{\alpha}$.)

If $\gamma_i \in [0,1], \forall i \in \overline{J}$, i.e. nodes are degraded, the new routing has two components:

1. Part of the total external arrival rate will be rejected, and
2. the admitted load will be redistributed among the nodes which are not completely broken down in a way to meet exactly the old ratios.

We will show that randomized reflection and skipping with acceptance probability vector $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\boldsymbol{\gamma})$ work, where $\alpha_0 = 1$ and $\alpha_i := \gamma_i$, $i \in \overline{J}$, constitute the vector
$\boldsymbol{\alpha}(\boldsymbol{\gamma}) = (\alpha_i, i \in \overline{J}_0)$ of acceptance probabilities as $\alpha_i := \gamma_i$.

If $\gamma_j \in [0,\infty)$, then we either speed up service at node j if $\gamma_j > 1$, or have a degraded server at node j if $\gamma_j < 1$. When at least one service rate increases, i.e. when $||\boldsymbol{\gamma}||_\infty > 1$, the mechanism to adapt the network's load and routing is:

- We increase the total network input by a factor $\beta = ||\boldsymbol{\gamma}||_\infty > 1$ to $\beta \cdot \lambda$ and
- redistribute the admitted load, choose with $\alpha_0 = 1$ as acceptance probability vector $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\boldsymbol{\gamma})$ the relative service rate changes $\alpha_j = \frac{\gamma_j}{||\boldsymbol{\gamma}||_\infty}, j \in \overline{J}$, for randomized reflection and skipping.

Remarks:

(i) If $\gamma_j > 1$, node j can process more load without being overloaded, which is easily seen by considering a single $M/M/1/\infty$ queue. In a network however, this additional load departing from j can cause overload at other nodes. Therefore some of the offered new total input of rate $\beta \cdot \lambda$ possibly will not be accepted after readjusting the routing. Our randomized random walk algorithms form Section 5 will automatically compute the rejection rates for the external arrivals at all nodes.

(ii) It was surprising to us that in the first case $\gamma_i \in [0,1], \forall i \in \overline{J}$ we have to choose exactly $\alpha_j = \gamma_j$ to adjust the acceptance probability of node j precisely to the service rate factor $\gamma_j \in [0,1]$. It means especially that our algorithm automatically detects the correct amount of load which is feasible to maintain the utilizations.

(iii) This observation lead us to conjecture the described $\boldsymbol{\alpha}$ in the general case.

The modified routing matrix is in both cases (randomized skipping, resp. reflection) denoted by $r^{(\boldsymbol{\alpha})} = (r^{(\boldsymbol{\alpha})}(i,j) : i,j \in \overline{J}_0)$, and is given alternatively by (5.1.7), resp. (5.2.1).

Definition 6.1.1. The set of "blocked nodes" $B(\boldsymbol{\gamma}) \subset \overline{J}$ is defined by $j \in B(\boldsymbol{\gamma}) :\Longleftrightarrow \gamma_j = 0$, and its complement in $\overline{J}$, the set of the "working nodes" $W(\boldsymbol{\gamma}) \subset \overline{J}$, is defined by $j \in W(\boldsymbol{\gamma}) :\Longleftrightarrow \gamma_j > 0$.

When the service rates of a Jackson network are modified according to $\boldsymbol{\gamma}$ and the routing is adjusted according to $\boldsymbol{\alpha}(\boldsymbol{\gamma})$ we obtain a new network process denoted by $\mathbf{X}^{(\gamma)} = (X^{(\gamma)}(t) : t \geq 0)$, the vector process recording the queue lengths in the network. $X_t^{(\gamma)} = (X_1^{(\gamma)}(t), \ldots, X_J^{(\gamma)}(t)) \in \mathbb{N}_0^{\overline{J}}$ reads: at time t there are $X_j^{(\gamma)}(t)$ customers present at node j, either in service or waiting. The assumptions put on the system imply that $\mathbf{X}$ is a strong Markov process on state space $\mathbb{N}_0^{\overline{J}}$ with generator $Q^{\mathbf{X}^{(\gamma)}} =: Q^{(\gamma)} = (q^{(\gamma)}(\mathbf{n}, \mathbf{n}') : \mathbf{n}, \mathbf{n}' \in \mathbb{N}_0^{\overline{J}})$. The strict positive transition rates of $Q^{(\gamma)}$ are under both rerouting regimes for $\mathbf{n} = (n_1, \ldots, n_J) \in \mathbb{N}_0^{\overline{J}}$

$$
\begin{aligned}
q^{(\gamma)}(\mathbf{n}, \mathbf{n} + \mathbf{e}_i) &= \beta\lambda r^{(\boldsymbol{\alpha})}(0,i), && i \in \overline{J}_0, \\
q^{(\gamma)}(\mathbf{n}, \mathbf{n} - \mathbf{e}_j + \mathbf{e}_i) &= 1_{[n_j>0]}\gamma_j\mu_j(n_j)r^{(\boldsymbol{\alpha})}(j,i), && i,j \in \overline{J},\, i \neq j, \\
q^{(\gamma)}(\mathbf{n}, \mathbf{n} - \mathbf{e}_j) &= 1_{[n_j>0]}\gamma_j\mu_j(n_j)r^{(\boldsymbol{\alpha})}(j,0), && j \in \overline{J}.
\end{aligned}
\tag{6.1.1}
$$

That the construction successfully in maintains the ratios (overall arrival rate/service rates) show the next theorems, which will be proved simultaneously. Recall $0/0 := 0$.

Theorem 6.1.2 (Modified Jackson networks: Change of service/rerouting)**.**

(a) [RANDOMIZED SKIPPING]*Let* $\mathbf{X}$ *be an ergodic Jackson network process as described in Section 6 with stationary distribution* ξ *from* (6.0.3)*, where the service intensities* $\mu_i(n_i)$ *at node* i *are changed by a factor* $\gamma_i \in [0, \infty)$ *for* $i \in \overline{J}$*. Denote*

$$\beta := \begin{cases} 1 & \text{if } ||\gamma||_\infty \leq 1, \\ ||\gamma||_\infty & \text{if } ||\gamma||_\infty > 1, \end{cases} \tag{6.1.2}$$

$$\alpha_0 = 1, \quad \text{and} \quad \alpha_j = \begin{cases} \gamma_j & \text{if } ||\boldsymbol{\gamma}||_\infty \leq 1, \\ \frac{\gamma_j}{||\boldsymbol{\gamma}||_\infty} & \text{if } ||\boldsymbol{\gamma}||_\infty > 1, \end{cases} \qquad \forall j \in \overline{J}, \tag{6.1.3}$$

and change routing by randomized skipping with $\boldsymbol{\alpha} = (\alpha_i : i \in \overline{J}_0)$ *according to Theorem 5.1.2, and change the total network input by factor* β*. Then* ξ *is a stationary distribution for* $\mathbf{X}^{(\boldsymbol{\gamma})} = (X^{(\boldsymbol{\gamma})}(t) : t \geq 0)$ *as well.*
Moreover, if $B(\boldsymbol{\gamma}) = \emptyset$*, then* $\mathbf{X}^{(\boldsymbol{\gamma})}$ *is ergodic.*
If $B(\boldsymbol{\gamma}) \neq \emptyset$*, then* $\mathbf{X}^{(\boldsymbol{\gamma})}$ *is not irreducible on* $\mathbb{N}_0^{\overline{J}}$ *and its state space is divided into an infinite set of closed subspaces*

$$\mathbb{N}_0^{W(\boldsymbol{\gamma})} \times \{(n_j : j \in B(\boldsymbol{\gamma}))\} \quad \forall (n_j : j \in B(\boldsymbol{\gamma})) \in \mathbb{N}_0^{B(\boldsymbol{\gamma})}.$$

For any probability distribution φ *on* $\mathbb{N}_0^{B(\boldsymbol{\gamma})}$ *there exists a stationary distribution* $\xi_\varphi^{(\boldsymbol{\gamma})}$ *for* $\mathbf{X}^{(\boldsymbol{\gamma})}$*, which is for* $\mathbf{n} = (n_1, \ldots, n_J) \in \mathbb{N}_0^{\overline{J}}$

$$\xi_\varphi^{(\boldsymbol{\gamma})}(\mathbf{n}) = \xi_\varphi^{(\boldsymbol{\gamma})}(n_1, \ldots, n_J) = \prod_{j \in W(\boldsymbol{\gamma})} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} C(j)^{-1} \cdot \varphi(n_j : j \in B(\boldsymbol{\gamma})). \tag{6.1.4}$$

(b) [RANDOMIZED REFLECTION]*If additionally to the assumptions of part (a) the routing chain* r *is reversible with respect to* η*, then the rerouting may be performed by randomized reflection according to Proposition 5.2.2 with* $\boldsymbol{\alpha} = (\alpha_i : i \in \overline{J}_0)$*. The results and formulas of part (a) carry over word by word.*

Proof. The global balance equation $x \cdot Q^{(\boldsymbol{\gamma})} = 0$ for the joint queue length process $X^{(\boldsymbol{\gamma})}$ of the modified system is in both settings for $\mathbf{n} = (n_1, \ldots, n_J) \in \mathbb{N}_0^{\overline{J}}$

$$\begin{aligned} &x(\mathbf{n}) \left(\sum_{j \in \overline{J}} \beta \lambda r^{(\boldsymbol{\alpha})}(0, j) + \sum_{j \in \overline{J}} 1_{[n_j > 0]} \gamma_j \mu_j(n_j)(1 - r^{(\boldsymbol{\alpha})}(j, j)) \right) \\ &= \sum_{i \in \overline{J}} x(\mathbf{n} - \mathbf{e}_i) 1_{[n_i > 0]} \beta \lambda r^{(\boldsymbol{\alpha})}(0, i) \\ &+ \sum_{j \in \overline{J}} \sum_{i \in \overline{J} \setminus \{j\}} x(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j) 1_{[n_i > 0]} \gamma_j \mu_j(n_j + 1) r^{(\boldsymbol{\alpha})}(j, i) \\ &+ \sum_{j \in \overline{J}} x(\mathbf{n} + \mathbf{e}_j) \gamma_j \mu_j(n_j + 1) r^{(\boldsymbol{\alpha})}(j, 0). \end{aligned} \tag{6.1.5}$$

We first consider the case $B(\boldsymbol{\gamma}) \neq \emptyset$. Then for $i \in B(\boldsymbol{\gamma})$ we have $\gamma_i = \alpha_i = 0$ and

$r^{(\boldsymbol{\alpha})}(j,i) = 0$ for all $j \in \overline{J}_0$, and (6.1.5) reduces to

$$\begin{aligned}
& x(\mathbf{n}) \left(\sum_{j \in W(\boldsymbol{\gamma})} \beta \lambda r^{(\boldsymbol{\alpha})}(0,j) + \sum_{j \in W(\boldsymbol{\gamma})} 1_{[n_j>0]} \gamma_j \mu_j(n_j)(1 - r^{(\boldsymbol{\alpha})}(j,j)) \right) \\
= & \sum_{i \in W(\boldsymbol{\gamma})} x(\mathbf{n} - \mathbf{e}_i) 1_{[n_i>0]} \beta \lambda r^{(\boldsymbol{\alpha})}(0,i) \\
& + \sum_{j \in W(\boldsymbol{\gamma})} \sum_{i \in W(\boldsymbol{\gamma}) \setminus \{j\}} x(\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j) 1_{[n_i>0]} \gamma_j \mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,i) \\
& + \sum_{j \in W(\boldsymbol{\gamma})} x(\mathbf{n} + \mathbf{e}_j) \gamma_j \mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,0).
\end{aligned} \tag{6.1.6}$$

Inserting $x(n_1,\ldots,n_J) = \prod_{j \in W(\boldsymbol{\gamma})} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} C(j)^{-1} \cdot \varphi(n_j : j \in B(\boldsymbol{\gamma}))$ for any probability density φ on $\mathbb{N}_0^{B(\boldsymbol{\gamma})}$ we see that immediately $\prod_{j \in W(\boldsymbol{\gamma})} C(j)^{-1} \cdot \varphi(n_j : j \in B(\boldsymbol{\gamma}))$ cancels. Multiplication with $\left(\beta \prod_{j \in W(\boldsymbol{\gamma})} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)}\right)^{-1}$ yields

$$\begin{aligned}
& \left(\sum_{j \in W(\boldsymbol{\gamma})} \lambda r^{(\boldsymbol{\alpha})}(0,j) + \sum_{j \in W(\boldsymbol{\gamma})} 1_{[n_j>0]} \frac{\gamma_j}{\beta} \mu_j(n_j)(1 - r^{(\boldsymbol{\alpha})}(j,j)) \right) \\
= & \sum_{i \in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]} \lambda r^{(\boldsymbol{\alpha})}(0,i) \\
& + \sum_{j \in W(\boldsymbol{\gamma})} \sum_{i \in W(\boldsymbol{\gamma}) \setminus \{j\}} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]} \frac{\eta_j}{\mu_j(n_j+1)} \frac{\gamma_j}{\beta} \mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,i) \\
& + \sum_{j \in W(\boldsymbol{\gamma})} \frac{\eta_j}{\mu_j(n_j+1)} \frac{\gamma_j}{\beta} \mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,0).
\end{aligned}$$

Using the fact that $\gamma_j/\beta = \gamma_j/||\boldsymbol{\gamma}||_\infty = \alpha_j$ for all $j \in \overline{J}$ we get the equation

$$\begin{aligned}
& \left(\sum_{j \in W(\boldsymbol{\gamma})} \lambda r^{(\boldsymbol{\alpha})}(0,j) + \sum_{j \in W(\boldsymbol{\gamma})} 1_{[n_j>0]} \alpha_j \mu_j(n_j)(1 - r^{(\boldsymbol{\alpha})}(j,j)) \right) \\
= & \sum_{i \in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]} \lambda r^{(\boldsymbol{\alpha})}(0,i) \\
& + \sum_{j \in W(\boldsymbol{\gamma})} \sum_{i \in W(\boldsymbol{\gamma}) \setminus \{j\}} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]} \frac{\eta_j}{\mu_j(n_j+1)} \alpha_j \mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,i) \\
& + \sum_{j \in W(\boldsymbol{\gamma})} \frac{\eta_j}{\mu_j(n_j+1)} \alpha_j \mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,0).
\end{aligned}$$

Reordering and canceling this yields

$$\begin{aligned}
&\left(\sum_{j\in W(\boldsymbol{\gamma})} \lambda r^{(\boldsymbol{\alpha})}(0,j) + \sum_{j\in W(\boldsymbol{\gamma})} 1_{[n_j>0]}\alpha_j\mu_j(n_j)\right)\\
&= \sum_{i\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\lambda r^{(\boldsymbol{\alpha})}(0,i)\\
&\quad + \sum_{j\in W(\boldsymbol{\gamma})}\sum_{i\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]} \frac{\eta_j}{\mu_j(n_j+1)}\alpha_j\mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,i)\\
&\quad + \sum_{j\in W(\boldsymbol{\gamma})} \frac{\eta_j}{\mu_j(n_j+1)}\alpha_j\mu_j(n_j+1) r^{(\boldsymbol{\alpha})}(j,0),
\end{aligned} \tag{6.1.7}$$

and so

$$\begin{aligned}
&\left(\sum_{j\in W(\boldsymbol{\gamma})} \lambda r^{(\boldsymbol{\alpha})}(0,j) + \sum_{j\in W(\boldsymbol{\gamma})} 1_{[n_j>0]}\alpha_j\mu_j(n_j)\right)\\
&= \sum_{i\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\lambda r^{(\boldsymbol{\alpha})}(0,i) + \sum_{j\in W(\boldsymbol{\gamma})}\sum_{i\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\alpha_j\eta_j r^{(\boldsymbol{\alpha})}(j,i)\\
&\quad + \sum_{j\in W(\boldsymbol{\gamma})} \eta_j\alpha_j r^{(\boldsymbol{\alpha})}(j,0).
\end{aligned} \tag{6.1.8}$$

The first term on the left side and the last term on the right side equate, because of

$$\sum_{j\in W(\boldsymbol{\gamma})} \lambda r^{(\boldsymbol{\alpha})}(0,j) = \lambda(1 - r^{(\boldsymbol{\alpha})}(0,0)), \quad \text{and}$$

$$\underbrace{\left(\sum_{j\in W(\boldsymbol{\gamma})} \eta_j\alpha_j r^{(\boldsymbol{\alpha})}(j,0) + \eta_0\alpha_0 r^{(\boldsymbol{\alpha})}(0,0)\right)}_{=\eta_0\alpha_0} - \eta_0\alpha_0 r^{(\boldsymbol{\alpha})}(0,0) = \lambda(1 - r^{(\boldsymbol{\alpha})}(0,0)),$$

where we used $\eta_0 = \lambda$, $\alpha_0 = 1$, and that $(\eta_j\alpha_j : j \in \overline{J}_0)$ is an invariant measure for $r^{(\boldsymbol{\alpha})}$. So (6.1.8) reduces to

$$\begin{aligned}
&\sum_{i\in W(\boldsymbol{\gamma})} 1_{[n_i>0]}\alpha_i\mu_i(n_i)\\
&= \sum_{i\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\lambda r^{(\boldsymbol{\alpha})}(0,i) + \sum_{j\in W(\boldsymbol{\gamma})}\sum_{i\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\alpha_j\eta_j r^{(\boldsymbol{\alpha})}(j,i).
\end{aligned} \tag{6.1.9}$$

Take any $i \in W(\boldsymbol{\gamma})$ with $n_i > 0$ and consider the summands with this i:

$$1_{[n_i>0]}\alpha_i\mu_i(n_i) = \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\lambda r^{(\boldsymbol{\alpha})}(0,i) + \sum_{j\in W(\boldsymbol{\gamma})} \frac{\mu_i(n_i)}{\eta_i} 1_{[n_i>0]}\alpha_j\eta_j r^{(\boldsymbol{\alpha})}(j,i),$$

which is

$$\alpha_i\eta_i \;=\; \lambda r^{(\boldsymbol{\alpha})}(0,i) + \sum_{j\in W(\boldsymbol{\gamma})} \alpha_j\eta_j r^{(\boldsymbol{\alpha})}(j,i),$$

and recalling $\eta_0 = \lambda$, $\alpha_0 = 1$ and $\alpha_j = 0$ for $j \in B(\boldsymbol{\gamma})$ this is

$$\alpha_i \eta_i = \sum_{j \in W(\boldsymbol{\gamma}) \cup \{0\}} \alpha_j \eta_j r^{(\boldsymbol{\alpha})}(j, i). \tag{6.1.10}$$

Note, that any i will occur in such a procedure for some suitable state vector with $n_i > 0$. Therefore, if (6.1.10) would be true, we would eventually arrive at

$$(\alpha_j \eta_j : j \in \overline{J}_0) \cdot r^{(\boldsymbol{\alpha})} = (\alpha_j \eta_j : j \in \overline{J}_0).$$

Now (6.1.10) is true for the setting of part (a) by Proposition 5.1.4 and $\alpha_0 = 1$, and for the setting of part (b) by the proof of Proposition 5.2.2 and $\alpha_0 = 1$, which finishes the parts for $B(\boldsymbol{\gamma}) \neq \emptyset$ of the proofs.

The case $B(\boldsymbol{\gamma}) = \emptyset$ is proved similarly. $\square$

Example 6.1.3. Consider a Jackson network on a node set $\overline{J} = \{1, 2, 3, 4\}$ with a routing matrix r as in Example 5.1.3 with service rate factors $\gamma = (1, 0.5, 1, 1)$. Then the availability vector is $\boldsymbol{\alpha} = (1, 1, 0.5, 1, 1)$ and a new modified routing matrix $r^{(\boldsymbol{\alpha})}$ is as in Example 5.1.3. See Figure 6.1.1 on page 169.

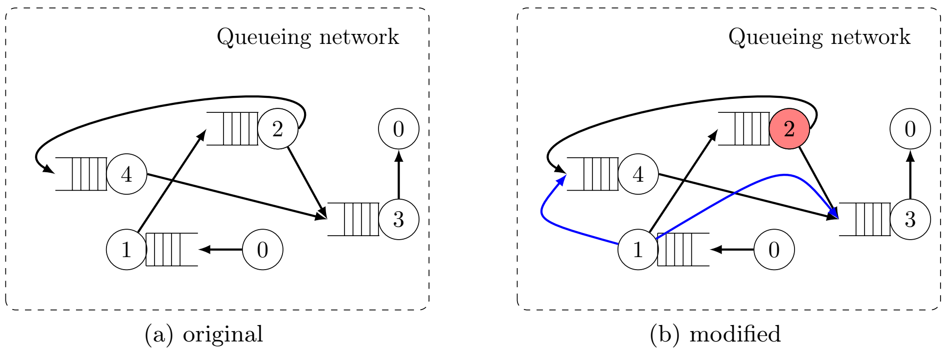


Figure 6.1.1.: Original and modified according to the skipping rule Jackson networks from Example 6.1.3.

Corollary 6.1.4. *If in the setting of Theorem 6.1.2 the Jackson network process* $\mathbf{X}$ *is ergodic with equilibrium* ξ *from* (6.0.3)*, and if after modification all nodes are still working, possibly with degraded capacity, i.e.* $B(\boldsymbol{\gamma}) = \emptyset$*, then in both cases of rerouting* $\mathbf{X}^{(\gamma)} = (X_t^{(\gamma)} : t \geq 0)$ *is ergodic with unique stationary and limiting distribution* ξ*.*

Corollary 6.1.5. *If in the framework of Theorem 6.1.2 we have* $r^{(\boldsymbol{\alpha})}(0,0) > 0$ *then the* effective arrival rate *after modification is* $\beta\lambda(1 - r^{(\boldsymbol{\alpha})}(0,0))$*.*

The following result summarizes the content of part **(a)** and **(b)** of Theorem 6.1.2 and extends both to an abstract framework.

Corollary 6.1.6 (Modified Jackson networks: Change of service, general rerouting)**.** *Let* $\mathbf{X}$ *be an ergodic Jackson network process as described in Section 6 with stationary distribution* ξ *from* (6.0.3)*, where the service intensities* $\mu_i(n_i)$ *at node* i *are changed by a factor* $\gamma_i \in [0, \infty)$ *for* $i \in \overline{J}$*. We change routing to follow some matrix* $r^{(\boldsymbol{\alpha})}$ *with invariant* *measure* $y = (\alpha_j \eta_j : j \in \overline{J}_0)$ *and increase the total network input by a factor* β*, where* $\alpha_0 = 1$*,* α_j *and* β *are defined as in* (6.1.3) *and* (6.1.2).

We denote the resulting Markovian state process on $\mathbb{N}_0^{\overline{J}}$ *by* $\mathbf{X}^{(\gamma)} = (X^{(\gamma)}(t) : t \geq 0)$.

Then ξ *is a stationary distribution for* $\mathbf{X}^{(\gamma)} = (X^{(\gamma)}(t) : t \geq 0)$ *as well.*

If $B(\boldsymbol{\gamma}) = \emptyset$*, then* $\mathbf{X}^{(\gamma)}$ *is ergodic.*

If $B(\boldsymbol{\gamma}) \neq \emptyset$*, then* $\mathbf{X}^{(\gamma)}$ *is not irreducible on* $\mathbb{N}_0^{\overline{J}}$ *and its state space is divided into an infinite set of closed subspaces* $\mathbb{N}_0^{W(\boldsymbol{\gamma})} \times \{(n_j : j \in B(\boldsymbol{\gamma}))\} \quad \forall (n_j : j \in B(\boldsymbol{\gamma})) \in \mathbb{N}_0^{B(\boldsymbol{\gamma})}$ *and for any probability distribution* φ *on* $\mathbb{N}_0^{B(\boldsymbol{\gamma})}$ *there exists a stationary distribution* $\xi_\varphi^{(\boldsymbol{\gamma})}$ *for* $\mathbf{X}^{(\gamma)}$ *given in* (6.1.4).

7. Jackson networks in a random environment

We consider a classical Jackson network from Section 6, the development of which is influenced by the time varying status of its external environment. On the other hand the network may trigger the environment to change its status. The dynamic is determined on one side by the environment as a jump process $\mathbf{Y} = (Y(t) : t \geq 0)$, changes of which result in changes of the network's parameter, and on the other side by the network process $\mathbf{X} = (X(t) : t \geq 0)$ as jump process where some jumps of $\mathbf{X}$ enforce the environment to immediately react to this jump. To be more precise:

The state (status) of the environment is recorded in a countable environment state space K and whenever the environment at time t is in state $Y(t) = k \in K$ it changes its status to $m \in K$ with rate $\nu(k, m)$.

The network process $\mathbf{X}$ records the joint queue length vector, as in Section 6, and $X_j(t) = n_j \in \mathbb{N}_0$ is the local queue length at node $j \in \overline{J}$. Whenever the environment is in state $k \in K$ and at node j a customer is served and leaves the network, then this jump of the local queue length triggers with probability $R_j(k, m)$ the environment to jump immediately from state k to $m \in K$.

Associated with environment state $k \in K$ is a vector $\boldsymbol{\gamma}(k) \in [0, \infty)^{\overline{J}}$ which determines the factor by which the service capacities are changed, when the environment enters k, similar to $\boldsymbol{\gamma} \in [0, \infty)^{\overline{J}}$ in Section 6.1. This results in a state dependent service rate $\mu_j(n_j, k) = \gamma_j(k) \cdot \mu_j(n_j)$ if the queue length at node j is n_j and the environment is k.

The network reacts to the impact of the environment in state k by modifying the routing according to different strategies, which we have described in Sections 5.1 and 5.2 and possibly with admitting more customers into the network. The latter part of the control strategy is set in force whenever in environment state k there exist some $\gamma_j(k) > 1$. In such state $k \in K$ the overall arrival rate to the network is increased by $\beta(\boldsymbol{\gamma}(k)) = ||\boldsymbol{\gamma}(k)||_\infty$ from λ to $\lambda \cdot \beta(\boldsymbol{\gamma}(k))$.

A schematic example of this kind of system is shown on Figure 7.0.1 on page 172.

We emphasize that neither the matrix $V = (v(k, m) : k, m \in K)$, which is (with suitable defined diagonal elements) a generator matrix nor the stochastic matrices $R_j = (R_j(k, m) : k, m \in K), j \in \overline{J}$, need to be irreducible or even ergodic, and furthermore that V is **not** the generator of $\mathbf{Y}$, which in general is not Markov.

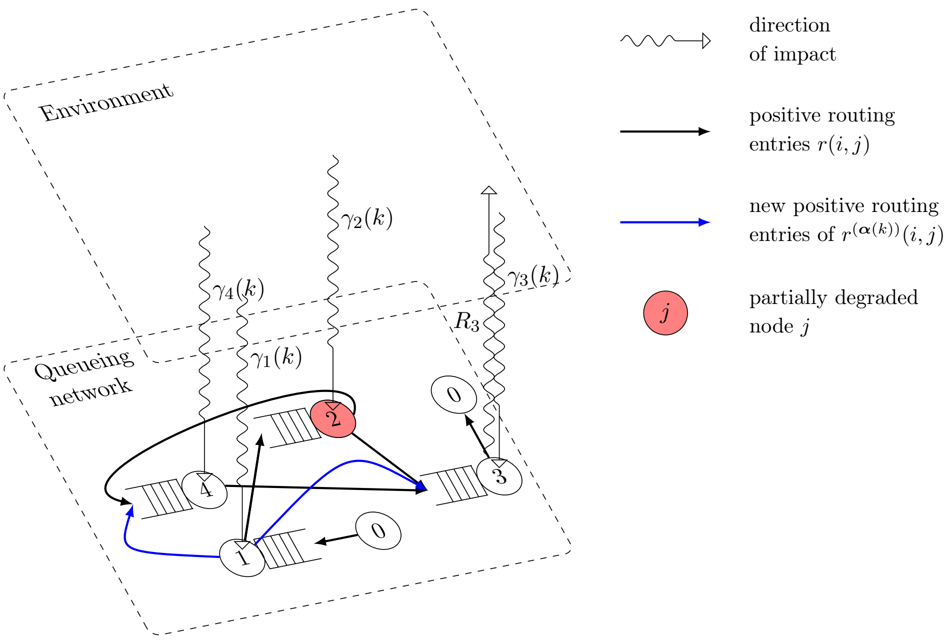


Figure 7.0.1.: Jackson network with degraded nodes – i.e. $\gamma_j(k) = \alpha_j(k) \leq 1$ for all j and k, and $\beta \equiv 1$ – in a random environment with two-way interaction. The service rates are modified by the environment. The routing is adopted according to the skipping rule. The customers, who leave the node 3, modify the environment.

7.1. Rerouting by randomized skipping

In this section we consider modification of routing in reaction to the servers' change of capacity by randomized skipping according to Section 5.1. We investigate this case in detail, while other modifications then can be described with less details.

We need environment dependent rerouting with acceptance probabilities $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\gamma(k))$, modified rerouting matrices $r^{(\boldsymbol{\alpha}(\gamma(k)))}$, and overall load factors $\beta(\gamma(k))$.

To keep notation short we write $\boldsymbol{\alpha}(k) = (\alpha_j(k) : j \in \overline{J}_0)$, instead of $\boldsymbol{\alpha}(\gamma(k))$, $r^{(\boldsymbol{\alpha}(k))}$ instead of $r^{(\boldsymbol{\alpha}(\gamma(k)))}$ and $\beta(k)$ instead of $\beta(\gamma(k))$.

The randomized skipping according to Section 5.1 yields a routing regime $r^{(\boldsymbol{\alpha}(k))}$ from Theorem 5.1.2, and the total service input rate is changed by a factor $\beta(k)$. $\boldsymbol{\alpha}$ and β are defined similar to (6.1.3) and (6.1.2), i.e. for $k \in K$:

$$\beta(k) := \begin{cases} 1 & \text{if} \quad ||\gamma(k)||_\infty \leq 1, \\ ||\gamma(k)||_\infty & \text{if} \quad ||\gamma(k)||_\infty > 1. \end{cases} \tag{7.1.1}$$

$$\alpha_0(k) = 1, \quad \text{and} \quad \alpha_j(k) = \begin{cases} \gamma_j(k) & \text{if} \quad ||\boldsymbol{\gamma}(k)||_\infty \leq 1, \\ \frac{\gamma_j(k)}{||\boldsymbol{\gamma}(k)||_\infty} & \text{if} \quad ||\boldsymbol{\gamma}(k)||_\infty > 1, \end{cases} \quad \forall j \in \overline{J}. \tag{7.1.2}$$

We further define $B(\boldsymbol{\gamma}(k))$ and $W(\boldsymbol{\gamma}(k))$ similar to Definition 6.1.1 as set of completely broken down nodes, resp. as set of nodes which, although possibly being degraded or upgraded, can still serve customers under environment condition k.

Definition 7.1.1. We denote the coupled process (QUEUE LENGTHS-ENVIRONMENT) by $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) = (Z(t) : t \geq 0) = ((X(t), Y(t)) : t \geq 0)$ on state space $E := \mathbb{N}_0^{\overline{J}} \times K$.

The dynamics of $\mathbf{Z}$ relies for the environment process $\mathbf{Y}$ on a generator matrix $V = (v(k,m) : k, m \in K)$ and stochastic matrices $R_j = (R_j(k,m) : k, m \in K), j \in \overline{J}$. Recall that the extended routing matrix $r = (r(i,j) : i, j \in \overline{J}_0)$ is irreducible and that $r^{(\boldsymbol{\alpha}(k))}$ from randomized skipping is irreducible on $W(\gamma(k)) \cup \{0\}$.

With the standard assumptions of independence for inter-arrival and service times and of conditional independence of routing and the jumps of the environment triggered by departing customers the queue lengths-environment process $\mathbf{Z}$ is a homogeneous Markov process on $E := \mathbb{N}_0^{\overline{J}} \times K$ with generator $Q^{\mathbf{Z}} = (q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n}', k')) : (\mathbf{n}, k), (\mathbf{n}', k') \in E)$. The strict positive transition rates of $Q^{\mathbf{Z}}$ are for $(\mathbf{n}, k) = ((n_1, \ldots, n_J), k) \in \mathbb{N}_0^{\overline{J}} \times K$ and $j, i \in \overline{J}$

$$\begin{aligned} q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n} + \mathbf{e}_i, k)) &= \beta(k)\lambda r^{(\boldsymbol{\alpha}(k))}(0, i), && (7.1.3) \\ q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n} - \mathbf{e}_j + \mathbf{e}_i, k)) &= 1_{[n_j > 0]}\gamma_j(k)\mu_j(n_j) r^{(\boldsymbol{\alpha}(k))}(j, i), && i \neq j, \\ q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n} - \mathbf{e}_j, m)) &= 1_{[n_j > 0]}\gamma_j(k)\mu_j(n_j) r^{(\boldsymbol{\alpha}(k))}(j, 0) R_j(k, m), && \\ q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n}, m)) &= v(k, m), && m \in K. \end{aligned}$$

Theorem 7.1.2. *Assume the queue lengths-environment process* $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ *from Definition 7.1.1 to be ergodic and assume that the pure Jackson network process* $\mathbf{X}$ *without environment is ergodic with stationary and limiting distribution* ξ *on* $\mathbb{N}_0^{\overline{J}}$ *from* (6.0.3)

$$\xi(\mathbf{n}) = \xi(n_1, \ldots, n_J) = \prod_{j=1}^{J} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} C(j)^{-1}, \quad \mathbf{n} \in \mathbb{N}_0^{\overline{J}},$$

with normalizing constants $C(j)$ *for the marginal (over nodes) distributions of* $\mathbf{X}$.

Define the reduced generator Q_{red} *as*

$$Q_{red} := \left[V + \sum_{j \in \overline{J}} \eta_j I_{(\gamma_j \bullet r^{(\boldsymbol{\alpha}(\cdot))}(j,0))}(R_j - I) \right], \tag{7.1.4}$$

where γ_j *and* $r^{(\boldsymbol{\alpha}(\cdot))}(j, 0)$ *are for* $j \in \overline{J}$ *real valued functions on* K. *Assume that the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a non zero, non-negative solution.*

Then Q_{red} *is irreducible on* K *and the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a strictly positive stochastic solution which we denote by* θ.

Furthermore, the queue lengths-environment process $\mathbf{Z}$ *has the unique steady-state distribution* $\pi = (\pi(\mathbf{n}, k) : \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K)$ *of product form:*

$$\pi(\mathbf{n}, k) = \xi(\mathbf{n})\theta(k), \quad \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, \ k \in K.$$

Remark: The reduced generator equation has a stochastic solution if $|K| < \infty$ holds, but there are many other easy to identify cases. The assumption, that $\theta \cdot Q_{\text{red}} = 0$ has a non zero, non-negative solution, rules out that Q_{red} is transient with only the zero vector as solution of that equation. As will be shown, the cases null-recurrent and transient with non-zero solution are not feasible due to ergodicity of $\mathbf{Z}$. This remark applies to Theorem 7.2.1 and Corollary 7.3.1 as well.

Proof. (of Theorem 7.1.2) The global balance equation of $\mathbf{Z}$ is

$$\pi(\mathbf{n},k)\Bigg(\sum_{i\in\overline{J}}\beta(k)\lambda r^{(\boldsymbol{\alpha}(k))}(0,i)+\underbrace{\sum_{m\in K\setminus\{k\}}v(k,m)}_{-\nu(k,k)}$$

$$+\sum_{j\in\overline{J}}1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)(1-r^{(\boldsymbol{\alpha}(k))}(j,j))\Bigg)$$

$$=\sum_{i\in\overline{J}}\pi(\mathbf{n}-\mathbf{e}_i,k)1_{[n_i>0]}\beta(k)\lambda r^{(\boldsymbol{\alpha}(k))}(0,i)$$

$$+\sum_{i\in\overline{J}}\sum_{j\in\overline{J}\setminus\{i\}}\pi(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j,k)1_{[n_i>0]}\gamma_j(k)\mu_j(n_j+1)r^{(\boldsymbol{\alpha}(k))}(j,i)$$

$$+\sum_{j\in\overline{J}}\sum_{m\in K}\pi(\mathbf{n}+\mathbf{e}_j,m)\gamma_j(m)\mu_j(n_j+1)r^{(\boldsymbol{\alpha}(m))}(j,0)R_j(m,k)$$

$$+\sum_{m\in K\setminus\{k\}}\pi(\mathbf{n},m)v(m,k).$$

Inserting $\pi(\mathbf{n},k)=\xi(\mathbf{n})\theta(k)$, adding $\xi(\mathbf{n})\theta(k)\cdot v(k,k)$ on both sides, and rearranging terms and blowing up this is

$$\theta(k)\left[\xi(\mathbf{n})\left(\sum_{i\in\overline{J}}\beta(k)\lambda r^{(\boldsymbol{\alpha}(k))}(0,i)+\sum_{j\in\overline{J}}1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)(1-r^{(\boldsymbol{\alpha}(k))}(j,j))\right)\right]$$

$$=\theta(k)\Bigg[\sum_{i\in\overline{J}}\xi(\mathbf{n}-\mathbf{e}_i)1_{[n_i>0]}\beta(k)\lambda r^{(\boldsymbol{\alpha}(k))}(0,i) \tag{7.1.5}$$

$$+\sum_{i\in\overline{J}}\sum_{j\in\overline{J}\setminus\{i\}}\xi(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j)1_{[n_i>0]}\gamma_j(k)\mu_j(n_j+1)r^{(\boldsymbol{\alpha}(k))}(j,i)$$

$$+\sum_{j\in\overline{J}}\xi(\mathbf{n}+\mathbf{e}_j)\gamma_j(k)\mu_j(n_j+1)r^{(\boldsymbol{\alpha}(k))}(j,0)\Bigg]$$

$$-\theta(k)\sum_{j\in\overline{J}}\xi(\mathbf{n}+\mathbf{e}_j)\gamma_j(k)\mu_j(n_j+1)r^{(\boldsymbol{\alpha}(k))}(j,0)$$

$$+\sum_{j\in\overline{J}}\sum_{m\in K}\xi(\mathbf{n}+\mathbf{e}_j)\theta(m)\gamma_j(m)\mu_j(n_j+1)r^{(\boldsymbol{\alpha}(m))}(j,0)R_j(m,k)$$

$$+\sum_{m\in K}\xi(\mathbf{n})\theta(m)v(m,k).$$

For each fixed environment state k the terms in squared brackets equate from Theorem 6.1.2, see (6.1.5), where for $B(\boldsymbol{\gamma}(k))$ we set in modified notation ($\varphi\to\varphi(k)$) from that theorem the specific probabilities

$$\varphi(k)(n_j:j\in B(\boldsymbol{\gamma}(k))):=\prod_{j\in B(\boldsymbol{\gamma}(k))}\prod_{\ell=1}^{n_j}\frac{\eta_j}{\mu_j(\ell)}C(j)^{-1},\quad (n_j:j\in B(\boldsymbol{\gamma}(k)))\in\mathbb{N}_0^{B(\boldsymbol{\gamma}(k))}.$$

Dividing by $\xi(\mathbf{n})$ and canceling $\mu_j(n_j+1)$ we arrive at

$$0 = -\theta(k)\sum_{j\in\overline{J}}\eta_j\gamma_j(k)r^{(\boldsymbol{\alpha}(k))}(j,0) + \sum_{j\in\overline{J}}\sum_{m\in K}\theta(m)\eta_j\gamma_j(m)r^{(\boldsymbol{\alpha}(m))}(j,0)R_j(m,k)$$
$$+\sum_{m\in K}\theta(m)v(m,k).$$

Rearranging terms we have

$$\theta(k)\sum_{j\in\overline{J}}\eta_j\gamma_j(k)r^{(\boldsymbol{\alpha}(k))}(j,0) = \sum_{m\in K}\theta(m)\left(v(m,k)+\sum_{j\in\overline{J}}\eta_j\gamma_j(m)r^{(\boldsymbol{\alpha}(m))}(j,0)R_j(m,k)\right),$$

which finally leads for any prescribed $k\in K$ to

$$0 = \sum_{m\in K}\theta(m)\left(v(m,k)+\sum_{j\in\overline{J}}\eta_j\gamma_j(m)r^{(\boldsymbol{\alpha}(m))}(j,0)(R_j(m,k)-\delta_{mk})\right). \tag{7.1.6}$$

This can be written in matrix form as

$$0 = \theta\underbrace{\left[V+\sum_{j\in\overline{J}}\eta_j I_{(\gamma_j\bullet r^{(\boldsymbol{\alpha}(\cdot))}(j,0))}(R_j-I)\right]}_{=:Q_{\text{red}}}. \tag{7.1.7}$$

So we have identified (7.1.4). Because the non diagonal elements of Q_{red} are non-negative whereas the row sum is zero, Q_{red} is the generator matrix of some Markov process.

Q_{red} is irreducible because otherwise $\mathbf{Z}$ would not be irreducible. By assumption, (7.1.7) has a non zero, non-negative solution. If the equation (7.1.7) has no stochastic solution the global balance equation of $\mathbf{Z}$ would have a non-trivial non-negative solution which cannot be normalized. This would contradict ergodicity. The same argument shows that the solution of (7.1.7) must be unique. □

Remark: Although the service rates $\mu_j(n_j)$ and routing probabilities $r(i,j)$ are locally determined with respect to the transition graph of r, the network control may in general be by global algorithms due to the applied randomized skipping by $r^{(\boldsymbol{\alpha})}$.

Careful analysis of the proof of Theorem 7.1.2 shows the following characterization.

Corollary 7.1.3. *Assume the queue lengths-environment process* $\mathbf{Z}=(\mathbf{X},\mathbf{Y})$ *from Definition 7.1.1 is irreducible. Define the* reduced generator *as*

$$Q_{red} := \left[V+\sum_{j\in\overline{J}}\eta_j I_{(\gamma_j\bullet r^{(\boldsymbol{\alpha}(\cdot))}(j,0))}(R_j-I)\right],$$

Then the following statements are equivalent:

(i) $\mathbf{Z}$ *is ergodic with product form steady-state distribution*

$$\pi(\mathbf{n}, k) = \xi(\mathbf{n}) \cdot \theta(k) = \prod_{j=1}^{J} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} C(j)^{-1} \cdot \theta(k), \qquad \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K.$$

(ii) *For all* $j \in \overline{J}$ *holds* $\sum_{n_j=0}^{\infty} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} < \infty$ *and the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a strictly positive stochastic solution.*

Remark: A similar corollary for Theorem 7.2.1 and Corollary 7.3.1 is obviously valid and will therefore not stated separately later on.

7.2. Rerouting by randomized reflection

In this section we assume that the modification of routing in reaction to the servers' change of capacities is by randomized reflection according to Section 5.2, which yields a routing regime $r^{(\boldsymbol{\alpha}(k))}$ according to Proposition 5.2.2. We use $\boldsymbol{\alpha}(k)$ and $\beta(k)$ as defined in (7.1.2) and (7.1.1), and take $B(\boldsymbol{\gamma}(k))$ and $W(\boldsymbol{\gamma}(k))$ as in Definition 6.1.1.

Recall, that the dynamics of the environment process $\mathbf{Y}$ is driven by a generator matrix $V = (\nu(k, m) : k, m \in K)$ and stochastic matrices $R_j = (R_j(k, m) : k, m \in K), j \in \overline{J}$, as described on p. 171. Note, that the original extended routing matrix $r = (r(i, j) : i, j \in \overline{J}_0)$ is irreducible but under randomized reflection $r^{(\boldsymbol{\alpha}(k))}$ may be reducible even on $W(\boldsymbol{\gamma}(k)) \cup \{0\}$, which does not destroy the ergodicity of the system process $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$. Then the queue lengths-environment process $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) = (Z(t) : t \geq 0) = ((X(t), Y(t)) : t \geq 0)$ is a homogeneous Markov process on state space $E := \mathbb{N}_0^{\overline{J}} \times K$ with generator $Q^{\mathbf{Z}} = (q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n}', k')) : (\mathbf{n}, k), (\mathbf{n}', k') \in E)$, which is formally identical to that displayed in (7.1.3).

As pointed out in Section 5.2 necessary for successfully applying randomized reflection as rerouting regime is reversibility of r, which we now set in force.

Theorem 7.2.1. *Consider the process* $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ *and assume that the extended routing matrix* $r = (r(i, j) : i, j \in \overline{J}_0)$ *is reversible for* $\eta = (\eta_j : j \in \overline{J}_0)$.

Assume $\mathbf{Z}$ *to be ergodic and assume that the pure Jackson network process* $\mathbf{X}$ *without environment is ergodic with stationary and limiting distribution* ξ *on* $\mathbb{N}_0^{\overline{J}}$ *from* (6.0.3)

$$\xi(\mathbf{n}) = \xi(n_1, \ldots, n_J) = \prod_{j=1}^{J} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} C(j)^{-1}, \quad \mathbf{n} \in \mathbb{N}_0^{\overline{J}}.$$

Define the reduced generator Q_{red} *as*

$$Q_{red} := \left[V + \sum_{j \in \overline{J}} \eta_j I_{(\gamma_j \bullet r^{(\boldsymbol{\alpha}(\cdot))}(j,0))} (R_j - I) \right],$$

where γ_j *and* $r^{(\boldsymbol{\alpha}(\cdot))}(j, 0)$ *are for* $j \in \overline{J}$ *real valued functions on* K*, and assume that the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a non zero, non-negative solution.*

Then Q_{red} *is irreducible on* K *and the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a strictly positive stochastic solution which we denote by* θ.

Furthermore, the queue lengths-environment process $\mathbf{Z}$ *has the unique steady-state distribution* $\pi = (\pi(\mathbf{n}, k) : \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K)$ *of product form:*

$$\pi(\mathbf{n}, k) = \xi(\mathbf{n})\theta(k), \qquad \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K.$$

The proof of the theorem is along the lines of the proof of Theorem 7.1.2, where we used almost completely the general abstract notation $r^{(\boldsymbol{\alpha}(k))}$ for the rerouting regime. Only when manipulating (7.1.5) we had to refer to properties of skipping in part (a) of Theorem 6.1.2, which is substituted now by referring to part (b) of that theorem.

Remark: The control of the customers' routing under randomized reflection by $r^{(\boldsymbol{\alpha})}$ is by local decisions with respect to the transition graph of r. So the network process is locally determined as well.

7.3. Rerouting by general randomization

The results of the previous sections suggests to extract general principles for randomized rerouting. We consider modifications of routing in reaction to servers' change of capacities by environment dependent factors $\boldsymbol{\gamma}(k) \in [0, \infty)^{\overline{J}}$ to $\mu_j(n_j, k) = \gamma_j(k)\mu_j(n_j)$. We use the notation introduced in Section 7.1 and $\boldsymbol{\alpha}(k)$ and $\beta(k)$ as defined in (7.1.2) and (7.1.1), and take $B(\boldsymbol{\gamma}(k))$ and $W(\boldsymbol{\gamma}(k))$ as in Definition 6.1.1.

The environment process $\mathbf{Y}$ is driven by $V = (v(k, m) : k, m \in K)$ and $R_j = (R_j(k, m) : k, m \in K), j \in \overline{J}$. For the general rerouting regimes $r^{(\boldsymbol{\alpha}(k))}, k \in K$, for $\boldsymbol{\alpha}(k)$ with $\alpha_0(k) = 1$ and $\boldsymbol{\alpha}(k) \in [0, \infty)^{\overline{J}_0}$, we only require the properties described in Corollary 6.1.6 and obtain a statement in the spirit of Zhu's main theorem [Zhu94, p. 12], and of Economou's Corollary 5 [Eco05, Section 5.2], where control regimes for rerouting are not specified. Our environment process is not Markov because of the two-way interaction of environment and service process, while Zhu's and Economou's theorem requires the environment to be Markov for its own.

Corollary 7.3.1. *The queue lengths-environment process* $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) = (Z(t) : t \geq 0) = ((X(t), Y(t)) : t \geq 0)$ *is a homogeneous Markov process on state space* $E := \mathbb{N}_0^{\overline{J}} \times K$ *with generator* $Q^{\mathbf{Z}} = (q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n}', k')) : (\mathbf{n}, k), (\mathbf{n}', k') \in E)$*, which is formally identical to that displayed in* (7.1.3)*.*

Assume that the rerouting regimes $r^{(\boldsymbol{\alpha}(k))}$*,* $k \in K$*, have invariant measures* $y(k) = (\alpha_j(k) \cdot \eta_j : j \in \overline{J}_0)$*.*

Assume $\mathbf{Z}$ *to be ergodic and assume that the pure Jackson network process* $\mathbf{X}$ *without environment is ergodic with stationary and limiting distribution* ξ *on* $\mathbb{N}_0^{\overline{J}}$ *from* (6.0.3)*.*

Define the reduced generator Q_{red} *as*

$$Q_{red} := \left[V + \sum_{j \in \overline{J}} \eta_j I_{(\gamma_j \bullet r^{(\boldsymbol{\alpha}(\cdot))}(j,0))}(R_j - I) \right],$$

where γ_j *and* $r^{(\boldsymbol{\alpha}(\cdot))}(j, 0)$ *are for* $j \in \overline{J}$ *real valued functions on* K*. Assume that the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a non zero, non-negative solution.*

Then Q_{red} *is irreducible on* K *and the* reduced generator equation $\theta \cdot Q_{red} = 0$ *has a strictly positive stochastic solution which we denote by* θ*.*

Furthermore, the queue lengths-environment process $\mathbf{Z}$ *has the unique steady-state distribution* $\pi = (\pi(\mathbf{n}, k) : \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K)$ *of product form:*

$$\pi(\mathbf{n}, k) = \xi(\mathbf{n})\theta(k), \quad \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K.$$

Remark: The environment dependent traffic equations $\eta^{(\boldsymbol{\alpha}(k))} = \eta^{(\boldsymbol{\alpha}(k))} \cdot r^{(\boldsymbol{\alpha}(k))}, k \in K$, have in general no strict positive solutions, because for $B(\boldsymbol{\gamma}(k)) \neq \emptyset$ and $j \in B(\boldsymbol{\gamma}(k))$ we have $\eta^{(\boldsymbol{\alpha}(k))}(j) = 0$. If there is some $k' \in K$ with $B(\boldsymbol{\gamma}(k')) = \emptyset$ we have $\eta^{(\boldsymbol{\alpha}(k'))}(j) > 0$. So the quotients $\eta^{(\boldsymbol{\alpha}(k))}(j)/\mu_j(n_j, k)$ are **not** independent of k, but nevertheless we obtain the product form steady-state distribution.

This observation should be compared with the first necessary condition and formula (3) in [Zhu94, Theorem 1] and with statements (i) and (ii) in [Eco05, Corollary 5], which prove equivalence of the existence of product form stationary distribution and invariance of the ratios (overall arrival rate/ service rates) under the condition that for all environment states the solution of the traffic equations are strictly positive [Eco14].

8. Environment changes dependent on queue lengths

In this section we start with the model from Section 7 for Jackson networks in a random environment and consider the situation where the generator V and the stochastic matrices R_j depend on the queuing state $\mathbf{n}$. We will analyze which additional conditions are sufficient to keep the steady-state distribution of the system in the similar form as in Corollary 7.3.1. The consequences of this change will be redefinition of measures ξ, θ and normalization constant and a weakening of the product form results.

Definition 8.0.2. We denote the coupled process (QUEUE LENGTHS-ENVIRONMENT) by $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) = (Z(t) : t \geq 0) = ((X(t), Y(t)) : t \geq 0)$ on state space $E := \mathbb{N}_0^{\overline{J}} \times K$.

The dynamics of $\mathbf{Z}$ relies for the environment process $\mathbf{Y}$ on generator matrices $V(\mathbf{n}) = (v(\mathbf{n}; k, m) : k, m \in K)$ and stochastic matrices $R_j(\mathbf{n}) = (R_j(\mathbf{n}; k, m) : k, m \in K)$, $j \in \overline{J}$, $\mathbf{n} \in \mathbb{N}_0^{\overline{J}}$.

Definition 8.0.3. For the extended routing matrices $r^{(k)}$ in Section 8 we only require that for all $k \in K$ there exists strictly positive vector $\boldsymbol{\eta} = (\eta_j : j \in \overline{J}_0)$ with $\eta_0 = \lambda$ and a fixed acceptance function $\boldsymbol{\alpha}(k) \in \{1\} \times [0,1]^{\overline{J}}$, such that it has an invariant measure $(\alpha_j(k)\eta_j : j \in \overline{J}_0)$

$$\alpha_j(k)\eta_j = \sum_{i=0}^{J} \alpha_i(k)\eta_i r^{(k)}(i,j), \qquad \forall j \in \overline{J}.$$

Example 8.0.4. It is now possible to have different extended rerouting matrices $r^{(k)}$ with the same invariant measure $(\alpha_j(k)\eta_j : j \in \overline{J}_0)$. Let $K = \{1,2\}$ and $\overline{J}_0 = \{0,1,2\}$, $\alpha_j(k) = 1$ for all $k \in K$ and all $j \in \overline{J}_0$,

$$r^{(1)} := \left(\begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & & 1 & \\ 1 & & & 1 \\ 2 & 1 & & \end{array}\right), \qquad r^{(2)} := \left(\begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & & & 1 \\ 1 & 1 & & \\ 2 & & 1 & \end{array}\right).$$

The system is a tandem network which switches the direction of the customer flow depending on the environment.

Definition 8.0.5. With the standard assumptions of independence for inter-arrival and service times and of conditional independence of routing and the jumps of the environment triggered by departing customers the queue lengths-environment process $\mathbf{Z}$ is a homogeneous Markov process on $E := \mathbb{N}_0^{\overline{J}} \times K$ with generator $Q^{\mathbf{Z}} = (q^{\mathbf{Z}}((\mathbf{n}, k), (\mathbf{n}', k')) : (\mathbf{n}, k), (\mathbf{n}', k') \in E)$. The non-negative transition rates of $Q^{\mathbf{Z}}$ are
for $(\mathbf{n}, k) = ((n_1, \ldots, n_J), k) \in \mathbb{N}_0^{\overline{J}} \times K$ and $i, j \in \overline{J}$

$$
\begin{aligned}
q^{\mathbf{Z}}((\mathbf{n},k),(\mathbf{n}+\mathbf{e}_i,k)) &= \beta(k)\lambda r^{(k)}(0,i), && \\
q^{\mathbf{Z}}((\mathbf{n},k),(\mathbf{n}-\mathbf{e}_j+\mathbf{e}_i,k)) &= 1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)r^{(k)}(j,i), && i\neq j, \\
q^{\mathbf{Z}}((\mathbf{n},k),(\mathbf{n}-\mathbf{e}_j,m)) &= 1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)r^{(k)}(j,0)R_j(\mathbf{n};k,m), && \\
q^{\mathbf{Z}}((\mathbf{n},k),(\mathbf{n},m)) &= v(\mathbf{n};k,m), && k\neq m.
\end{aligned}
\tag{8.0.1}
$$

Here we use the same α, β, γ, λ, μ, $\overline{J}$ as in Section 7. The rerouting matrices $r^{(k)}$ are general rerouting matrices defined in Definition 8.0.3:

$$
\gamma_j(k) \in [0,\infty),
$$

$$
\beta(k) := \begin{cases} 1 & \text{if } ||\gamma(k)||_\infty \leq 1, \\ ||\gamma(k)||_\infty & \text{if } ||\gamma(k)||_\infty > 1, \end{cases}
$$

$$
\alpha_0(k) = 1, \quad \text{and} \quad \alpha_j(k) = \begin{cases} \gamma_j(k) & \text{if } |\boldsymbol{\gamma}(k)||_\infty \leq 1, \\ \frac{\gamma_j(k)}{||\boldsymbol{\gamma}(k)||_\infty} & \text{if } ||\boldsymbol{\gamma}(k)||_\infty > 1, \end{cases}
$$

for all $k \in K$ and $j \in \overline{J}$. The general rerouting regimes $r^{(k)}$ have unique invariant measure $(\alpha_j(k)\eta_j : j \in \overline{J}_0)$ for all $k \in K$ and some strictly positive $(\eta_j : j \in \overline{J}_0)$ as specified in Definition 8.0.3.

Definition 8.0.6. Similar to Section 7.1 we define for each $k \in K$ the set of "blocked nodes" $B(\boldsymbol{\gamma}(k)) \subset \overline{J}$. It is defined by $j \in B(\boldsymbol{\gamma}(k)) :\Longleftrightarrow \gamma_j(k) = 0$. Its complement in $\overline{J}$, the set of the "working nodes" $W(\boldsymbol{\gamma}(k)) \subset \overline{J}$, is defined by $j \in W(\boldsymbol{\gamma}(k)) :\Longleftrightarrow \gamma_j(k) > 0$.

Before we present the main theorem, we need results for modified Jackson networks similar to Corollary 6.1.6. In contrast to Corollary 6.1.6, we only focus on one particular, not normalized and not necessary unique solution ξ of the equation $\xi Q^{(\boldsymbol{\gamma})} = 0$.

Lemma 8.0.7. *Given generator* $Q^{(\boldsymbol{\gamma})} = (q^{(\boldsymbol{\gamma})}(\mathbf{n},\mathbf{n}') : \mathbf{n},\mathbf{n}' \in \mathbb{N}_0^{\overline{J}})$

$$
\begin{aligned}
q^{(\boldsymbol{\gamma})}(\mathbf{n},\mathbf{n}+\mathbf{e}_i) &= \beta\lambda r_G(0,i), && i \in \overline{J}_0, \\
q^{(\boldsymbol{\gamma})}(\mathbf{n},\mathbf{n}-\mathbf{e}_j+\mathbf{e}_i) &= 1_{[n_j>0]}\gamma_j\mu_j(n_j)r_G(j,i), && i,j \in \overline{J},\ i\neq j, \\
q^{(\boldsymbol{\gamma})}(\mathbf{n},\mathbf{n}-\mathbf{e}_j) &= 1_{[n_j>0]}\gamma_j\mu_j(n_j)r_G(j,0), && j \in \overline{J},
\end{aligned}
\tag{8.0.2}
$$

with $\gamma_j \in [0,\infty)$

$$
\beta := \begin{cases} 1 & \text{if } ||\gamma||_\infty \leq 1 \\ ||\gamma||_\infty & \text{if } ||\gamma||_\infty > 1, \end{cases}
$$

$$
\alpha_0 = 1, \quad \text{and} \quad \alpha_j = \begin{cases} \gamma_j & \text{if } ||\boldsymbol{\gamma}||_\infty \leq 1, \\ \frac{\gamma_j}{||\boldsymbol{\gamma}||_\infty} & \text{if } ||\boldsymbol{\gamma}||_\infty > 1, \end{cases} \qquad \forall j \in \overline{J},
$$

any stochastic routing matrix r_G *with invariant measure* $y = (\alpha_j\eta_j : j \in \overline{J}_0)$, $\alpha_0 = \lambda$, $\eta_0 = \lambda$ and $\eta_j > 0$ for all $j \in \overline{J}$. *Then the measure*

$$
\xi(\mathbf{n}) := \xi(n_1,\ldots,n_J) = \prod_{j\in\overline{J}}\prod_{\ell=1}^{n_j}\frac{\eta_j}{\mu_j(\ell)} \tag{8.0.3}
$$

solves the global balance equation

$$
\xi Q^{(\boldsymbol{\gamma})} = 0.
$$

Proof. The proof is similar to the proof of Theorem 6.1.2 except the routing matrix r_{G} can be any stochastic matrix with invariant measure $y = (\alpha_j \eta_j : j \in \overline{J}_0)$ and, to keep the notational burden small, we focus only on one particular and not normalized measure ξ defined in (8.0.3).

The global balance equation $\xi \cdot Q^{(\gamma)} = 0$ for $\mathbf{n} = (n_1, \ldots, n_J) \in \mathbb{N}_0^{\overline{J}}$ is

$$\begin{aligned}
&\xi(\mathbf{n}) \left(\sum_{j\in\overline{J}} \beta\lambda r_{\mathrm{G}}(0,j) + \sum_{j\in\overline{J}} 1_{[n_j>0]}\gamma_j\mu_j(n_j)(1-r_{\mathrm{G}}(j,j)) \right) \\
&= \sum_{i\in\overline{J}} \xi(\mathbf{n}-\mathbf{e}_i)1_{[n_i>0]}\beta\lambda r_{\mathrm{G}}(0,i) \\
&\quad + \sum_{j\in\overline{J}} \sum_{i\in\overline{J}\setminus\{j\}} \xi(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j)1_{[n_i>0]}\gamma_j\mu_j(n_j+1)r_{\mathrm{G}}(j,i) \\
&\quad + \sum_{j\in\overline{J}} \xi(\mathbf{n}+\mathbf{e}_j)\gamma_j\mu_j(n_j+1)r_{\mathrm{G}}(j,0).
\end{aligned} \tag{8.0.4}$$

Inserting $\xi(n_1,\ldots,n_J) = \prod_{j\in\overline{J}}\prod_{\ell=1}^{n_j}\frac{\eta_j}{\mu_j(\ell)}$ and multiplication with $\left(\beta\prod_{j\in\overline{J}}\prod_{\ell=1}^{n_j}\frac{\eta_j}{\mu_j(\ell)}\right)^{-1}$ yields

$$\begin{aligned}
&\left(\sum_{j\in\overline{J}} \lambda r_{\mathrm{G}}(0,j) + \sum_{j\in\overline{J}} 1_{[n_j>0]}\frac{\gamma_j}{\beta}\mu_j(n_j)(1-r_{\mathrm{G}}(j,j)) \right) \\
&= \sum_{i\in\overline{J}} \frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\lambda r_{\mathrm{G}}(0,i) \\
&\quad + \sum_{j\in\overline{J}} \sum_{i\in\overline{J}\setminus\{j\}} \frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\frac{\eta_j}{\mu_j(n_j+1)}\frac{\gamma_j}{\beta}\mu_j(n_j+1)r_{\mathrm{G}}(j,i) \\
&\quad + \sum_{j\in\overline{J}} \frac{\eta_j}{\mu_j(n_j+1)}\frac{\gamma_j}{\beta}\mu_j(n_j+1)r_{\mathrm{G}}(j,0).
\end{aligned}$$

Using the fact that $\gamma_j/\beta = \begin{cases} \gamma_j/1 = \alpha_j, & \gamma_j \le 1 \\ \gamma_j/||\boldsymbol{\gamma}||_\infty = \alpha_j, & \gamma_j > 1 \end{cases}$ for all $j \in \overline{J}$ we get the equation

$$\begin{aligned}
&\left(\sum_{j\in\overline{J}} \lambda r_{\mathrm{G}}(0,j) + \sum_{j\in\overline{J}} 1_{[n_j>0]}\alpha_j\mu_j(n_j)(1-r_{\mathrm{G}}(j,j)) \right) \\
&= \sum_{i\in\overline{J}} \frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\lambda r_{\mathrm{G}}(0,i) + \sum_{j\in\overline{J}} \sum_{i\in\overline{J}\setminus\{j\}} \frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\frac{\eta_j}{\mu_j(n_j+1)}\alpha_j\mu_j(n_j+1)r_{\mathrm{G}}(j,i) \\
&\quad + \sum_{j\in\overline{J}} \frac{\eta_j}{\mu_j(n_j+1)}\alpha_j\mu_j(n_j+1)r_{\mathrm{G}}(j,0).
\end{aligned}$$

Reordering and canceling this yields

$$
\begin{aligned}
&\left(\sum_{j\in\overline{J}}\lambda r_{\mathrm{G}}(0,j)+\sum_{j\in\overline{J}}1_{[n_j>0]}\alpha_j\mu_j(n_j)\right)\\
&=\sum_{i\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\lambda r_{\mathrm{G}}(0,i)\\
&\quad+\sum_{j\in\overline{J}}\sum_{i\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\frac{\eta_j}{\mu_j(n_j+1)}\alpha_j\mu_j(n_j+1)r_{\mathrm{G}}(j,i)\\
&\quad+\sum_{j\in\overline{J}}\frac{\eta_j}{\mu_j(n_j+1)}\alpha_j\mu_j(n_j+1)r_{\mathrm{G}}(j,0),
\end{aligned}
$$

and so

$$
\begin{aligned}
&\left(\sum_{j\in\overline{J}}\lambda r_{\mathrm{G}}(0,j)+\sum_{j\in\overline{J}}1_{[n_j>0]}\alpha_j\mu_j(n_j)\right)\\
&=\sum_{i\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\lambda r_{\mathrm{G}}(0,i)+\sum_{j\in\overline{J}}\sum_{i\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\alpha_j\eta_j r_{\mathrm{G}}(j,i)\\
&\quad+\sum_{j\in\overline{J}}\eta_j\alpha_j r_{\mathrm{G}}(j,0).
\end{aligned}
\tag{8.0.5}
$$

The first term on the left side and the last term on the right side equate, because of

$$
\begin{aligned}
&\sum_{j\in\overline{J}}\lambda r_{\mathrm{G}}(0,j)=\lambda(1-r_{\mathrm{G}}(0,0)),\quad\text{and}\\
&\underbrace{\left(\sum_{j\in\overline{J}}\eta_j\alpha_j r_{\mathrm{G}}(j,0)+\eta_0\alpha_0 r_{\mathrm{G}}(0,0)\right)}_{=\eta_0\alpha_0}-\eta_0\alpha_0 r_{\mathrm{G}}(0,0)=\lambda(1-r_{\mathrm{G}}(0,0)),
\end{aligned}
$$

where we used $\eta_0=\lambda$, $\alpha_0=1$, and that $(\eta_j\alpha_j : j\in\overline{J}_0)$ is an invariant measure for r_{G}. So (8.0.5) reduces to

$$
\begin{aligned}
&\sum_{i\in\overline{J}}1_{[n_i>0]}\alpha_i\mu_i(n_i)\\
&=\sum_{i\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\lambda r_{\mathrm{G}}(0,i)+\sum_{j\in\overline{J}}\sum_{i\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}1_{[n_i>0]}\alpha_j\eta_j r_{\mathrm{G}}(j,i)
\end{aligned}
\tag{8.0.6}
$$

$$
\begin{aligned}
&\iff 0=\sum_{i\in\overline{J}}1_{[n_i>0]}\left(\frac{\mu_i(n_i)}{\eta_i}\underbrace{\lambda}_{=\eta_0\alpha_0}r_{\mathrm{G}}(0,i)+\sum_{j\in\overline{J}}\frac{\mu_i(n_i)}{\eta_i}\alpha_j\eta_j r_{\mathrm{G}}(j,i)-\alpha_i\mu_i(n_i)\right)\\
&\iff 0=\sum_{i\in\overline{J}}1_{[n_i>0]}\frac{\mu_i(n_i)}{\eta_i}\left(\sum_{j\in\overline{J}_0}\alpha_j\eta_j r_{\mathrm{G}}(j,i)-\alpha_i\eta_i\right).
\end{aligned}
$$

The difference in the parenthesis is 0 because $(\alpha_j\eta_j : j \in \overline{J}_0)$ is invariant measure of r_{G}. $\square$

Lemma 8.0.8. *Let r_G be some rerouting matrix from Lemma 8.0.7, then*

$$r_G(j,i) = 0, \text{ if } j \in W(\boldsymbol{\gamma}) \cup \{0\} \text{ and } i \in B(\boldsymbol{\gamma})$$

with $W(\boldsymbol{\gamma})$ and $B(\boldsymbol{\gamma})$ from Definition 6.1.1.

Proof. Recall that all η_j are positive. Let i be any node from $B(\boldsymbol{\gamma})$, then $\eta_i\alpha_i = 0$. The products $\eta_j\alpha_j$ are positive for all $j \in W(\boldsymbol{\gamma})\cup\{0\}$ and they are 0 for all $j \in B(\boldsymbol{\gamma})$. Because $(\alpha_j\eta_j : j \in \overline{J}_0)$ is invariant measure of r_{G} we have for $i \in B(\boldsymbol{\gamma})$

$$\sum_{j\in\overline{J}_0} \eta_j\alpha_j r_{\mathrm{G}}(j,i) = \eta_i\alpha_i = 0 \qquad \Longleftrightarrow \qquad \sum_{j\in W(\boldsymbol{\gamma})\cup\{0\}} \underbrace{\eta_j\alpha_j}_{>0} r_{\mathrm{G}}(j,i) = 0.$$

$\square$

Theorem 8.0.9. *Assume the queue lengths-environment process $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ from Definition 8.0.5 to be ergodic. Furthermore assume that there is non-zero and non-negative solution $\theta = (\theta(\mathbf{n}; k) : \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, k \in K) \in \mathbb{R}^{\overline{J}\times K}$ of the equations*

$$\theta(\mathbf{n})Q_{red}(\mathbf{n}) = 0, \qquad \forall \mathbf{n} \in \mathbb{N}_0^{\overline{J}},$$

with

$$\theta(\mathbf{n}) = (\theta(\mathbf{n}; k) : k \in K),$$

and

$$Q_{red}(\mathbf{n}) := \left[V(\mathbf{n}) + \sum_{j\in\overline{J}} \eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}(R_j(\mathbf{n}+\mathbf{e}_j) - I) \right], \tag{8.0.7}$$

such that

$$\theta(\mathbf{n}; k) = \theta(\mathbf{n} + \mathbf{e}_j; k), \qquad \forall k \in K, \mathbf{n} \in \mathbb{N}_0^{\overline{J}}, j \in W(\boldsymbol{\gamma}(k)). \tag{8.0.8}$$

Then the steady-state solution of equation $\pi Q = 0$ is

$$\pi(\mathbf{n}, k) = C^{-1}\xi(\mathbf{n})\theta(\mathbf{n}; k), \tag{8.0.9}$$

with

$$\xi(\mathbf{n}) := \prod_{j=1}^{J}\prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)}$$

and the normalization constant

$$C = \sum_{\mathbf{n}\in\mathbb{N}_0^{\overline{J}}} \sum_{k\in K} \xi(\mathbf{n})\theta(\mathbf{n}, k) < \infty. \tag{8.0.10}$$

Recall, by convention on page 149: For each $j \in \overline{J}$ is $I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}$ a diagonal matrix from $\mathbb{R}^{K\times K}$ with entries

$$I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}(k,k) = \gamma_j(k) r^{(k)}(j,0), \qquad \forall k \in K,$$
$$I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}(k,m) = 0, \qquad \forall k \neq m.$$

Remark 8.0.10. The expression (8.0.9) is not a true product form. Any stochastic measure $\pi(\mathbf{n},k)$ can be expressed as $\xi(\mathbf{n})\theta(\mathbf{n};k)$ by setting $\theta(\mathbf{n};k) = \pi(\mathbf{n},k)/\xi(\mathbf{n})$ for each k and $\mathbf{n}$.

Remark 8.0.11. The equation (8.0.8) means that $\theta(\mathbf{n};k)$ and $\theta(\mathbf{n}';k)$ must be equal if $\mathbf{n}$ differs from $\mathbf{n}'$ in at least one working node j. We can also write this as

$$\forall k \in K, \forall \mathbf{n}, \mathbf{n}' \in \mathbb{N}_0^{\overline{J}} : (\exists_{j \in \overline{J}} : n_j \neq n_j^{'} \wedge \gamma_j(k) \neq 0) \Longrightarrow \theta(\mathbf{n};k) = \theta(\mathbf{n}';k)$$

or, equivalently, for any k and all $\mathbf{n}$ the vector $\theta(\mathbf{n};k)$ is equal to some $\tilde{\theta}(\mathbf{n}|_{B(\boldsymbol{\gamma}(k))}, k)$ which only depends on the set of blocked nodes $B(\boldsymbol{\gamma}(k))$.

Proof of Theorem 8.0.9. The global balance equations of $\mathbf{Z}$ are for $(\mathbf{n};k) \in E$

$$\begin{aligned}
&\pi(\mathbf{n},k)\Bigg(\sum_{i\in\overline{J}} \beta(k)\lambda r^{(k)}(0,i) + \underbrace{\sum_{m\in K\setminus\{k\}} v(\mathbf{n};k,m)}_{-v(\mathbf{n};k,k)} \\
&\qquad\qquad + \sum_{j\in\overline{J}} 1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)(1 - r^{(k)}(j,j))\Bigg) \\
&= \sum_{i\in\overline{J}} \pi(\mathbf{n}-\mathbf{e}_i,k)1_{[n_i>0]}\beta(k)\lambda r^{(k)}(0,i) \\
&\quad + \sum_{i\in\overline{J}}\sum_{j\in\overline{J}\setminus\{i\}} \pi(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j,k)1_{[n_i>0]}\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,i) \\
&\quad + \sum_{j\in\overline{J}}\sum_{m\in K} \pi(\mathbf{n}+\mathbf{e}_j,m)\gamma_j(m)\mu_j(n_j+1)r^{(m)}(j,0)R_j(\mathbf{n}+\mathbf{e}_j;m,k) \\
&\quad + \sum_{m\in K\setminus\{k\}} \pi(\mathbf{n},m)v(\mathbf{n};m,k).
\end{aligned} \tag{8.0.11}$$

Inserting $\pi(\mathbf{n},k) = C^{-1}\xi(\mathbf{n})\theta(\mathbf{n};k)$, canceling C^{-1}, adding $\xi(\mathbf{n})\theta(\mathbf{n};k)\cdot v(\mathbf{n};k,k)$ on both sides, and rearranging terms and blowing up leads to

$$\begin{aligned}
&\xi(\mathbf{n})\theta(\mathbf{n};k)\left(\sum_{i\in\overline{J}}\beta(k)\lambda r^{(k)}(0,i) + \sum_{j\in\overline{J}} 1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)(1-r^{(k)}(j,j))\right)\\
&= \sum_{i\in\overline{J}}\xi(\mathbf{n}-\mathbf{e}_i)\theta(\mathbf{n}-\mathbf{e}_i;k)1_{[n_i>0]}\beta(k)\lambda r^{(k)}(0,i)\\
&\quad+\sum_{i\in\overline{J}}\sum_{j\in\overline{J}\setminus\{i\}}\xi(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j)\theta(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j;k)1_{[n_i>0]}\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,i)\\
&\quad+\left\{\sum_{j\in\overline{J}}\xi(\mathbf{n}+\mathbf{e}_j)\theta(\mathbf{n}+\mathbf{e}_j;k)\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,0)\right.\\
&\quad\left.-\sum_{j\in\overline{J}}\xi(\mathbf{n}+\mathbf{e}_j)\theta(\mathbf{n}+\mathbf{e}_j;k)\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,0)\right\}\\
&\quad+\sum_{j\in\overline{J}}\sum_{m\in K}\xi(\mathbf{n}+\mathbf{e}_j)\theta(\mathbf{n}+\mathbf{e}_j;m)\gamma_j(m)\mu_j(n_j+1)r^{(m)}(j,0)R_j(\mathbf{n}+\mathbf{e}_j;m,k)\\
&\quad+\sum_{m\in K}\xi(\mathbf{n})\theta(\mathbf{n};m)v(\mathbf{n};m,k).
\end{aligned}\tag{8.0.12}$$

According to Lemma 8.0.8 it holds for any $k\in K$

$$\begin{aligned}
\theta(\mathbf{n}-\mathbf{e}_i;k)r^{(k)}(0,i) = 0 &= \theta(\mathbf{n};k)r^{(k)}(0,i), &&\forall i\in B(\gamma(k)),\\
\theta(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j;k)\gamma_j(k)r^{(k)}(j,i) = 0 &= \theta(\mathbf{n};k)\gamma_j(k)r^{(k)}(j,i), &&\forall j\in W(\gamma(k))\\
&&&\wedge\, i\in B(\gamma(k)).
\end{aligned}$$

From definition of $\gamma_j(k)=0 \iff j\in B(\boldsymbol{\gamma}(k))$ follows

$$\begin{aligned}
\theta(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j;k)\gamma_j(k)r^{(k)}(j,i) = 0 &= \theta(\mathbf{n};k)\gamma_j(k)r^{(k)}(j,i), &&\forall j\in B(\boldsymbol{\gamma}(k)),\\
\theta(\mathbf{n}+\mathbf{e}_j;k)\gamma_j(k) = 0 &= \theta(\mathbf{n};k)\gamma_j(k), &&\forall j\in B(\boldsymbol{\gamma}(k)).
\end{aligned}$$

From assumption (8.0.8) it follows

$$\begin{aligned}
\theta(\mathbf{n}-\mathbf{e}_i;k) &= \theta(\mathbf{n};k), &&\forall i\in W(\boldsymbol{\gamma}(k)),\\
\theta(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j;k) &= \theta(\mathbf{n};k), &&\forall i,j\in W(\boldsymbol{\gamma}(k)),\\
\theta(\mathbf{n}+\mathbf{e}_j;k) &= \theta(\mathbf{n};m), &&\forall j\in W(\boldsymbol{\gamma}(k)).
\end{aligned}$$

Therefore we can replace all $\theta(\mathbf{n}-\mathbf{e}_i;k)$, $\theta(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j;k)$, $\theta(\mathbf{n}+\mathbf{e}_j;k)$ in the global balance equations (8.0.12) by $\theta(\mathbf{n};k)$ and obtain

$$
\begin{aligned}
&\theta(\mathbf{n};k)\left[\xi(\mathbf{n})\Big(\sum_{i\in\overline{J}}\beta(k)\lambda r^{(k)}(0,i)+\sum_{j\in\overline{J}}1_{[n_j>0]}\gamma_j(k)\mu_j(n_j)(1-r^{(k)}(j,j))\Big)\right]\\
&=\theta(\mathbf{n};k)\Bigg[\sum_{i\in\overline{J}}\xi(\mathbf{n}-\mathbf{e}_i)1_{[n_i>0]}\beta(k)\lambda r^{(k)}(0,i)\\
&\quad+\sum_{i\in\overline{J}}\sum_{j\in\overline{J}\setminus\{i\}}\xi(\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j)1_{[n_i>0]}\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,i)\\
&\quad+\sum_{j\in\overline{J}}\xi(\mathbf{n}+\mathbf{e}_j)\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,0)\Bigg]\\
&\quad-\theta(\mathbf{n};k)\sum_{j\in\overline{J}}\xi(\mathbf{n}+\mathbf{e}_j)\gamma_j(k)\mu_j(n_j+1)r^{(k)}(j,0)\\
&\quad+\sum_{j\in\overline{J}}\sum_{m\in K}\theta(\mathbf{n};m)\xi(\mathbf{n}+\mathbf{e}_j)\gamma_j(m)\mu_j(n_j+1)r^{(m)}(j,0)R_j(\mathbf{n}+\mathbf{e}_j;m,k)\\
&\quad+\sum_{m\in K}\xi(\mathbf{n})\theta(\mathbf{n},m)v(\mathbf{n};m,k).
\end{aligned}
\tag{8.0.13}
$$

In the same way as in the proof of Theorem 7.1.2 for each fixed environment state k the terms in squared brackets are formulas (8.0.4) from Lemma 8.0.7, therefore they are equal for the proposed measure $\xi(\mathbf{n})=\prod_{j=1}^{J}\prod_{\ell=1}^{n_j}\frac{\eta_j}{\mu_j(\ell)}$ and we can cancel them.

Dividing by $\xi(\mathbf{n})$ and canceling $\mu_j(n_j+1)$ we arrive at

$$
\begin{aligned}
0=&-\theta(\mathbf{n};k)\sum_{j\in\overline{J}}\eta_j\gamma_j(k)r^{(k)}(j,0)\\
&+\sum_{j\in\overline{J}}\sum_{m\in K}\theta(\mathbf{n};m)\eta_j\gamma_j(m)r^{(m)}(j,0)R_j(\mathbf{n}+\mathbf{e}_j;m,k)\\
&+\sum_{m\in K}\theta(\mathbf{n};m)v(\mathbf{n};m,k).
\end{aligned}
$$

Rearranging terms we have

$$
\begin{aligned}
&\theta(\mathbf{n};k)\sum_{j\in\overline{J}}\eta_j\gamma_j(k)r^{(k)}(j,0)\\
&=\sum_{m\in K}\theta(\mathbf{n};m)\left(v(\mathbf{n};m,k)+\sum_{j\in\overline{J}}\eta_j\gamma_j(m)r^{(m)}(j,0)R_j(\mathbf{n}+\mathbf{e}_j;m,k)\right),
\end{aligned}
$$

which finally leads for any prescribed $k\in K$ to

$$
0=\sum_{m\in K}\theta(\mathbf{n};m)\left(v(\mathbf{n};m,k)+\sum_{j\in\overline{J}}\eta_j\gamma_j(m)r^{(m)}(j,0)(R_j(\mathbf{n}+\mathbf{e}_j;m,k)-\delta_{mk})\right).
\tag{8.0.14}
$$

This can be written in matrix form as

$$0 = \theta(\mathbf{n}) \underbrace{\left[V(\mathbf{n}) + \sum_{j \in \overline{J}} \eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))} (R_j(\mathbf{n} + \mathbf{e}_j) - I) \right]}_{=Q_{\text{red}}(\mathbf{n})}. \tag{8.0.15}$$

So we have identified (8.0.7).

With $\xi(\mathbf{n})\theta(\mathbf{n}; k)$ we found a non-zero and non-negative solution of the global balance equations (8.0.11) of the process $\mathbf{Z}$. Because $\mathbf{Z}$ is ergodic the normalization constant C defined in (8.0.10) exists and $\pi(\mathbf{n}, k) = C^{-1}\xi(\mathbf{n})\theta(\mathbf{n}; k)$ is the unique steady-state distribution of the system. □

Corollary 8.0.12. *For any* $\mathbf{n}$ *the* $Q_{red}(\mathbf{n})$ *in Theorem 8.0.9 is a generator.*

Proof. $Q_{\text{red}}(\mathbf{n})$ is defined as

$$Q_{\text{red}}(\mathbf{n}) := V(\mathbf{n}) + \sum_{j \in \overline{J}} \eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))} (R_j(\mathbf{n} + \mathbf{e}_j) - I). \tag{8.0.16}$$

For any $j \in \overline{J}$ and any $\mathbf{n} \in \mathbb{N}_0^{\overline{J}}$ matrix $\eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}(R_j(\mathbf{n} + \mathbf{e}_j) - I)$ can have negative values only in its diagonal. Because all matrices $R_j(\mathbf{n} + \mathbf{e}_j)$ are stochastic the row sum of the resulting matrix is 0:

$$\left(\eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))} (R_j(\mathbf{n} + \mathbf{e}_j) - I) \right) \mathbf{e} = \left(\eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))} (\mathbf{e} - \mathbf{e}) \right) = 0.$$

Therefore each matrix $\eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}(R_j(\mathbf{n}+\mathbf{e}_j)-I)$ is a generator. Consequently for any $\mathbf{n} \in \mathbb{N}_0^{\overline{J}}$ the matrix $Q_{\text{red}}(\mathbf{n})$, which is a sum of generators $V(\mathbf{n})$ and $\eta_j I_{(\gamma_j \bullet r^{(\cdot)}(j,0))}(R_j(\mathbf{n} + \mathbf{e}_j) - I)$, is the generator matrix of some Markov process. □

In the following we will give some examples of systems with non-constant $V(\mathbf{n})$ and $R_j(\mathbf{n})$. They are constructed in such a way that the solutions $\theta(\mathbf{n})$ of equations $\theta(\mathbf{n})Q_{\text{red}}(\mathbf{n}) = 0$ with $Q_{\text{red}}(\mathbf{n})$ defined by (8.0.7) have property (8.0.8).

Definition 8.0.13. We define a "masking" function Δ_B which replaces n_j in $\mathbf{n}$ by 0 if node j is NOT blocked

$$\Delta_B(\mathbf{n}; k) := \begin{cases} n_j, & \gamma_j(k) = 0 \iff j \in B(\boldsymbol{\gamma}(k)), \\ 0, & \gamma_j(k) \neq 0 \iff j \notin B(\boldsymbol{\gamma}(k)). \end{cases}$$

Note, that Δ_B implicitly depends on $\boldsymbol{\gamma}$. Because the dependence will be clear from the context, we omit this to simplify notation.

From the definition it follows

$$\Delta_B(\mathbf{n}; k) = \Delta_B(\mathbf{n}'; k) \iff \forall j \in B(\boldsymbol{\gamma}(k)) : n_j = n_j'$$

and

$$\Delta_B(\mathbf{n}; k) = \Delta_B(\mathbf{n} + \mathbf{e}_j; k), \qquad \forall j \in W(\boldsymbol{\gamma}(k)). \tag{8.0.17}$$

Proposition 8.0.14. *Given an environment set $K = \{1, 2\}$ and functions*

$$\begin{aligned} A &: \mathbb{N}_0^{\overline{J}} \times K \times K \to \mathbb{R}_0^+, \\ B &: \mathbb{N}_0^{\overline{J}} \times K \times K \to \mathbb{R}^+, \\ C &: \overline{J} \times \mathbb{N}_0^{\overline{J}} \times K \times K \to \mathbb{R}_0^+, \end{aligned}$$

with

$$C(j, \mathbf{n}, k, m) \leq B(\mathbf{n}, k, m), \qquad \forall j, \mathbf{n}, k, m.$$

The entries of generators $V(\mathbf{n})$ in Theorem 8.0.9 are set to

$$v(\mathbf{n}; k, m) := \frac{A(\Delta_B(\mathbf{n}; m), k, m)}{B(\Delta_B(\mathbf{n}; k), k, m)} \qquad \text{for } k \neq m.$$

The entries of stochastic matrices $R_j(\mathbf{n})$ are set to

$$\begin{aligned} R_j(\mathbf{n} + \mathbf{e}_j; k, m) &:= \frac{C(j, \Delta_B(\mathbf{n} + \mathbf{e}_j; m), k, m)}{B(\Delta_B(\mathbf{n}; k), k, m)} \qquad \text{for } k \neq m. \\ R_j(\mathbf{n} + \mathbf{e}_j; k, k) &:= 1 - R_j(\mathbf{n} + \mathbf{e}_j; k, m) \end{aligned}$$

Then θ with

$$\begin{aligned} \theta(\mathbf{n}, 1) = & \left(A(\Delta_B(\mathbf{n}; 1), 2, 1) + \sum_{j \in \overline{J}} \eta_j \gamma_j(2) r^{(2)}(j, 0) C(j, \Delta_B(\mathbf{n} + \mathbf{e}_j; 1), 2, 1) \right) \\ & \cdot B(\Delta_B(\mathbf{n}; 1), 1, 2), \\ \theta(\mathbf{n}, 2) = & \left(A(\Delta_B(\mathbf{n}; 2), 1, 2) + \sum_{j \in \overline{J}} \eta_j \gamma_j(1) r^{(1)}(j, 0) C(j, \Delta_B(\mathbf{n} + \mathbf{e}_j; 2), 1, 2) \right) \\ & \cdot B(\Delta_B(\mathbf{n}; 2), 2, 1), \end{aligned}$$

solves the equations $\theta(\mathbf{n}) Q_{red}(\mathbf{n}) = 0$ (8.0.7) and θ has property (8.0.8).

Proof. According to (8.0.7), the matrices $Q_{\text{red}}(\mathbf{n}) \in \mathbb{R}^{K \times K}$ are then

$$Q_{\text{red}}(\mathbf{n}) = \begin{pmatrix} -Q_{\text{red}}(\mathbf{n})_{12} & Q_{\text{red}}(\mathbf{n})_{12} \\ Q_{\text{red}}(\mathbf{n})_{21} & -Q_{\text{red}}(\mathbf{n})_{21} \end{pmatrix}$$

with

$$\begin{aligned} Q_{\text{red}}(\mathbf{n})_{12} &= v(\mathbf{n}; 1, 2) + \sum_{j \in \overline{J}} \eta_j \gamma_j(1) r^{(1)}(j, 0) R_j(\mathbf{n} + \mathbf{e}_j, 1, 2) \\ &= \frac{A(\Delta_B(\mathbf{n}; 2)_j, 1, 2) + \sum_{j \in \overline{J}} \eta_j \gamma_j(1) r^{(1)}(j, 0) C(\Delta_B(\mathbf{n} + \mathbf{e}_j; 2), 1, 2)}{B(\Delta_B(\mathbf{n}; 1), 1, 2)} \\ &=: \frac{a(\mathbf{n}, 1, 2)}{b(\mathbf{n}, 1, 2)} \end{aligned}$$

and

$$\begin{aligned}
Q_{\text{red}}(\mathbf{n})_{21} &= v(\mathbf{n};2,1) + \sum_{j\in\overline{J}} \eta_j\gamma_j(2) r^{(2)}(j,0) R_j(\mathbf{n}+\mathbf{e}_j,2,1) \\
&= \frac{A(\Delta_B(\mathbf{n};1),2,1) + \sum_{j\in\overline{J}} \eta_j\gamma_j(2) r^{(2)}(j,0) C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j;1),2,1)}{B(\Delta_B(\mathbf{n};2),2,1)} \\
&=: \frac{a(\mathbf{n},2,1)}{b(\mathbf{n},2,1)}.
\end{aligned}$$

For each $\mathbf{n}$ matrix $Q_{\text{red}}(\mathbf{n})$ has a form

$$Q_{\text{red}}(\mathbf{n}) = \begin{pmatrix} -\frac{a(\mathbf{n},1,2)}{b(\mathbf{n},1,2)} & \frac{a(\mathbf{n},1,2)}{b(\mathbf{n},1,2)} \\ \frac{a(\mathbf{n},2,1)}{b(\mathbf{n},2,1)} & -\frac{a(\mathbf{n},2,1)}{b(\mathbf{n},2,1)} \end{pmatrix}.$$

and vector

$$\theta(\mathbf{n}) := (a(\mathbf{n},2,1)\cdot b(\mathbf{n},1,2), a(\mathbf{n},1,2)\cdot b(\mathbf{n},2,1))$$

solves $\theta(\mathbf{n})Q_{\text{red}}(\mathbf{n}) = 0$.

Using definition of $a(\mathbf{n},2,1)$ and $b(\mathbf{n},2,1)$ we get

$$\begin{aligned}
\theta(\mathbf{n},1) &= \left(A(\Delta_B(\mathbf{n};1),2,1) + \sum_{j\in\overline{J}} \eta_j\gamma_j(2) r^{(2)}(j,0) C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j;1),2,1) \right) \\
&\quad \cdot B(\Delta_B(\mathbf{n};1),1,2), \\
\theta(\mathbf{n},2) &= \left(A(\Delta_B(\mathbf{n};2),1,2) + \sum_{j\in\overline{J}} \eta_j\gamma_j(1) r^{(1)}(j,0) C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j;2),1,2) \right) \\
&\quad \cdot B(\Delta_B(\mathbf{n};2),2,1).
\end{aligned}$$

We show that θ has property (8.0.8). We write a detailed proof for entries $\theta(\mathbf{n},1)$. The proof for $\theta(\mathbf{n},2)$ is the same, but with 1 replaced by 2 and vice versa.

Let $i \in W(\gamma(1))$ then it holds

$$\begin{aligned}
&\theta(\mathbf{n}+\mathbf{e}_i,1) \\
&= \left(A(\Delta_B(\mathbf{n}+\mathbf{e}_i;1),2,1) + \sum_{j\in\overline{J}} \eta_j\gamma_j(2) r^{(2)}(j,0) C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j+\mathbf{e}_i;1),2,1) \right) \\
&\quad \cdot B(\Delta_B(\mathbf{n}+\mathbf{e}_i;1),1,2).
\end{aligned} \tag{8.0.18}$$

Due to property (8.0.17) of the masking function Δ_B, we have for any $i \in W(\boldsymbol{\gamma}(1))$:

$$\begin{aligned}
\Delta_B(\mathbf{n}+\mathbf{e}_i;1) &= \Delta_B(\mathbf{n};1) \\
\text{and } \Delta_B(\mathbf{n}+\mathbf{e}_j+\mathbf{e}_i;1) &= \Delta_B(\mathbf{n}+\mathbf{e}_j;1).
\end{aligned} \tag{8.0.19}$$

Substituting (8.0.19) into (8.0.18) yields $\theta(\mathbf{n}+\mathbf{e}_i,1) = \theta(\mathbf{n},1)$ for any $i \in W(\gamma_i(1))$, which is property (8.0.8). $\square$

In the following Definition 8.0.15 we will model a Jackson network where a failure of a single node immediately leads to freezing of the whole network system. Then an exponentially distributed repair process depending only on the number of customers in the system is started. As soon as the repair is finished, the work of the whole network in unfrozen and the work is resumed.

Because failure (repair) and freezing (unfreezing) of the whole system happen simultaneously and the repair process does not need to "know" which node causes the failure, we can only focus on freezing and unfreezing states of the network.

Definition 8.0.15 (Stalling)**.** Given a Jackson network in a random environment from Definition 8.0.5. The environment space $K := \{\emptyset, \overline{J}\}$ describes the set of the down nodes. Only two environment states are possible

- $\emptyset \in K$ if all nodes are up, the system is unfrozen.
- $\overline{J} \in K$ if all nodes are down , the system is frozen.

From definition of states in K, it follows $\boldsymbol{\gamma}(\emptyset) := (1, \ldots, 1)$ and $\gamma(\overline{J}) := (0, \ldots, 0)$. We assume value $\beta \equiv 1$ to be constant.

We assume the rerouting matrix $r^{(\emptyset)}$– when all nodes are up – to be irreducible. And we require that the rerouting matrix is $r^{(\overline{J})}$ – when all nodes are down – is:

$$r^{(\overline{J})} := \left(\begin{array}{c|cc} & 0 & (j : j \in \overline{J}) \\ \hline 0 & 1 & 0 \\ (\overline{J} : j \in \overline{J}) & \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} & \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \end{array} \right) = I. \tag{8.0.20}$$

Remark 8.0.16. We can choose the rerouting matrix $r^{(\overline{J})}$ rather general:

$$r^{(\overline{J})} = \left(\begin{array}{c|cc} & 0 & (j : j \in \overline{J}) \\ \hline 0 & 1 & 0 \\ (\overline{J} : j \in \overline{J}) & M_0 & M_{\overline{J}} \end{array} \right)$$

with some stochastic matrix $(M_0|M_{\overline{J}}) \in [0,1]^{\overline{J}\times\overline{J}_0}$. The important property of $r^{(\overline{J})}$ is that $(\lambda, 0, \ldots, 0)$ is an invariant measure of $r^{(\overline{J})}$. For our stalling model we have chosen the simplest version of $r^{(\overline{J})}$, given in (8.0.20).

Special case of the system from Definition 8.0.15 with $R_j(\mathbf{n}) = I$, are Jackson networks with stalling in described in [SD03a, Section 5.2]. That means, in [SD03a, Section 5.2], the customers cannot enforce a breakdown of the network on departure.

In the following Corollary 8.0.17 we will use Proposition 8.0.14 to extend the model of a degradable Jackson networks controlled by Stalling-principle from [SD03a, Section 5.2]. Using non-trivial matrices $R_j(\mathbf{n})$ we will add the possibility to break down for the system each times a customer leaves the system.

Corollary 8.0.17. *Consider an ergodic degradable Jackson network on a node set $\overline{J}$ from Definition 8.0.15.*

The failure rates, which are independent from the arrival and service process, are $\nu_{failure}(\mathbf{n}) \in \mathbb{R}_0^+$, the repair rates are $\nu_{repair}(\mathbf{n}) \in \mathbb{R}^+$. The queueing system fails with the *probability $p_{failure}(j,\mathbf{n}) \in [0,1]$ at a time instant when customer leaves the network from the node j and there were $\mathbf{n}$ customers in the system right before.*

Then the steady-state distribution $\pi(\mathbf{n},k)$ is

$$\begin{aligned}\pi(\mathbf{n},\emptyset) &= C^{-1}\xi(\mathbf{n}),\\ \pi(\mathbf{n},\overline{J}) &= C^{-1}\xi(\mathbf{n})\frac{\nu_{failure}(\mathbf{n}) + \sum_{j\in\overline{J}}\eta_j r^{(\emptyset)}(j,0)p_{failure}(j,\mathbf{n}+\mathbf{e}_j)}{\nu_{repair}(\mathbf{n})},\end{aligned} \tag{8.0.21}$$

with the normalization constant

$$C = \sum_{\mathbf{n}\in\mathbb{N}_0^{\overline{J}}}\xi(\mathbf{n})\left(1+\frac{\nu_{failure}(\mathbf{n}) + \sum_{j\in\overline{J}}\eta_j r^{(\emptyset)}(j,0)p_{failure}(j,\mathbf{n}+\mathbf{e}_j)}{\nu_{repair}(\mathbf{n})}\right) < \infty. \tag{8.0.22}$$

Remark 8.0.18. The ν_{repair} of the ergodic system must be positive. Assume there is some $\mathbf{n}'$ such that $v_{repair}(\mathbf{n}')$ is 0, then the system may never leave the state $(\mathbf{n}',\overline{J})$ with all blocked nodes. The absorbing state $(\mathbf{n}',\overline{J})$ contradicts the ergodicity assumption.

If all $p_{failure}(j,\mathbf{n}) = 0$ then we have a system from [SD03a, Theorem 5.5]. In this case also the rates $\nu_{failure}$ must be positive. Assume there is some $\mathbf{n}'$ such that $\nu_{failure}(\mathbf{n}') = 0$ then the system will never reach the inessential state $(\mathbf{n}',\overline{J})$ again when it has departed from its communicating class.

Proof of Corollary 8.0.17. With $\boldsymbol{\gamma}(\emptyset) = (1,\dots,1)$, $\boldsymbol{\gamma}(\overline{J}) = (0,\dots,0)$, and $\beta = 1$ it holds

$$\begin{aligned}\boldsymbol{\alpha}(\emptyset) &= \boldsymbol{\gamma}(\emptyset) = (1,\dots,1),\\ \boldsymbol{\alpha}(\overline{J}) &= \boldsymbol{\gamma}(\overline{J}) = (0,\dots,0).\end{aligned}$$

We show that rerouting matrix $r^{(\cdot)}$ satisfies routing properties from Definition 8.0.3. We need to determine the positive vector $\boldsymbol{\eta} = (\eta_j : j \in \overline{J}_0)$ with $\eta_0 = \lambda$ and show that $(\alpha_j(k)\eta_j : j \in \overline{J}_0)$ is an invariant measure of $r^{(k)}$ for $k \in \{\emptyset, \overline{J}\}$.

The stochastic matrix $r^{(\emptyset)}$ is irreducible, therefore there exists a unique positive invariant measure $\boldsymbol{\eta}$ with $\boldsymbol{\eta} r^{(\emptyset)} = \boldsymbol{\eta}$ and $\eta_0 = \lambda$. Obviously the vector $(\alpha_j(\emptyset)\eta_j : j \in \overline{J}_0) = \boldsymbol{\eta}$ is an invariant measure of $r^{(\emptyset)}$. The vector $(\alpha_j(\overline{J})\eta_j : j \in \overline{J}_0) = (\lambda,0,\dots,0)$ is an invariant measure of the matrix $r^{(\overline{J})} = I$.

Following Proposition 8.0.14 with $1\hat{=}\emptyset$ and $2\hat{=}\overline{J}$ we define

$$\begin{aligned}A(\mathbf{n},\emptyset,\overline{J}) &= \nu_{failure}(\mathbf{n}), & A(\mathbf{n},\overline{J},\emptyset) &= 1,\\ B(\mathbf{n},\emptyset,\overline{J}) &= 1, & B(\mathbf{n},\overline{J},\emptyset) &= \nu_{repair}(\mathbf{n})^{-1},\end{aligned}$$

$$\begin{aligned}C(j,\mathbf{n},\emptyset,\overline{J}) &= p_{failure}(j,\mathbf{n}),\\ C(j,\mathbf{n},\overline{J},\emptyset) &= 0.\end{aligned}$$

It holds

$$\begin{aligned}v(\mathbf{n};\emptyset,\overline{J}) &:= \frac{A(\Delta_B(\mathbf{n};\overline{J}),\emptyset,\overline{J})}{B(\Delta_B(\mathbf{n};\emptyset),\emptyset,\overline{J})} = \frac{A(\mathbf{n},\emptyset,\overline{J})}{B(\mathbf{0},\emptyset,\overline{J})} = \frac{\nu_{failure}(\mathbf{n})}{1},\\ v(\mathbf{n};\overline{J},\emptyset) &:= \frac{A(\Delta_B(\mathbf{n};\emptyset),\overline{J},\emptyset)}{B(\Delta_B(\mathbf{n};\overline{J}),\overline{J},\emptyset)} = \frac{A(\mathbf{0},\overline{J},\emptyset)}{B(\mathbf{n},\overline{J},\emptyset)} = \frac{1}{\nu_{repair}(\mathbf{n})^{-1}}.\end{aligned}$$

$$V(\mathbf{n}) = \begin{array}{c|cc} & \emptyset & \overline{J} \\ \hline \emptyset & -\nu_{failure}(\mathbf{n}) & \nu_{failure}(\mathbf{n}) \\ \overline{J} & \nu_{repair}(\mathbf{n}) & -\nu_{repair}(\mathbf{n}) \end{array} .$$

$$R_j(\mathbf{n}+\mathbf{e}_j;\emptyset,\overline{J}) := \frac{C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j;\overline{J}),\emptyset,\overline{J})}{B(\Delta_B(\mathbf{n};\emptyset),\emptyset,\overline{J})} = \frac{C(j,\mathbf{n}+\mathbf{e}_j,\emptyset,\overline{J})}{B(\mathbf{0},\emptyset,\overline{J})} = \frac{p_{failure}(j,\mathbf{n}+\mathbf{e}_j)}{1},$$
$$R_j(\mathbf{n}+\mathbf{e}_j;\overline{J},\emptyset) := \frac{C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j;\emptyset),\overline{J},\emptyset)}{B(\Delta_B(\mathbf{n};\overline{J}),\overline{J},\emptyset)} = \frac{C(j,\mathbf{0},\overline{J},\emptyset)}{B(\mathbf{n},\overline{J},\emptyset)} = \frac{0}{1} = 0.$$

$$R_j(\mathbf{n}) = \begin{array}{c|cc} & \emptyset & \overline{J} \\ \hline \emptyset & 1-p_{failure}(j,\mathbf{n}) & p_{failure}(j,\mathbf{n}) \\ \overline{J} & 0 & 1 \end{array} .$$

$$\begin{aligned} \theta(\mathbf{n};\emptyset) = & \left(A(\Delta_B(\mathbf{n};\emptyset),\overline{J},\emptyset) + \sum_{j\in\overline{J}} \eta_j \underbrace{\gamma_j(\overline{J})}_{=0} r^{(\overline{J})}(j,0) C(j,\Delta_B(\mathbf{n}+\mathbf{e}_j;\emptyset),\overline{J},\emptyset) \right) \\ & \cdot B(\Delta_B(\mathbf{n};\emptyset),\emptyset,\overline{J}) \\ = & A(\mathbf{0},\overline{J},\emptyset)B(\mathbf{0},\emptyset,\overline{J}) = 1, \end{aligned} \tag{8.0.23}$$

$$\begin{aligned} \theta(\mathbf{n};\overline{J}) = & \left(A(\underbrace{\Delta_B(\mathbf{n};\overline{J})}_{=\mathbf{n}},\emptyset,\overline{J}) + \sum_{j\in\overline{J}} \eta_j \underbrace{\gamma_j(\emptyset)}_{=1} r^{(\emptyset)}(j,0) C(j,\underbrace{\Delta_B(\mathbf{n}+\mathbf{e}_j;\overline{J})}_{=\mathbf{n}+\mathbf{e}_j},\emptyset,\overline{J}) \right) \\ & \cdot B(\underbrace{\Delta_B(\mathbf{n};\overline{J})}_{=\mathbf{n}},\overline{J},\emptyset) \\ = & \frac{\nu_{failure}(\mathbf{n}) + \sum_{j\in\overline{J}} \eta_j r^{(\emptyset)}(j,0) p_{failure}(j,\mathbf{n}+\mathbf{e}_j)}{\nu_{repair}(\mathbf{n})}. \end{aligned} \tag{8.0.24}$$

The solution θ is not negative and not zero: In (8.0.23) and (8.0.24) we see that all entries of vector θ are not negative. From (8.0.23) we conclude that at least one entry of θ, namely $\theta(\mathbf{n};\emptyset)$, is not zero.

The normalization constant is

$$C = \sum_{\mathbf{n}\in\mathbb{N}_0^{\overline{J}}} \sum_{k\in\{\emptyset,\overline{J}\}} \xi(\mathbf{n})\theta(\mathbf{n};k) = \sum_{\mathbf{n}\in\mathbb{N}_0^{\overline{J}}} \xi(\mathbf{n}) \left(\theta(\mathbf{n};\emptyset) + \theta(\mathbf{n};\overline{J})\right).$$

The unique steady-state distribution (8.0.21) follows from Theorem 8.0.9. □

Corollary 8.0.19 ([SD03a, Theorem 6.4])**.** *Consider an ergodic degradable Jackson-Network on a node set $\overline{J}$ as defined in Definition 8.0.5. The environment space $K = 2^{\overline{J}}$ describes the sets of the down nodes. The rerouting matrix $r^{(\emptyset)}$, when all nodes are up, is* *irreducible. The rerouting matrix $r^{(k)}$ is $r^{(\boldsymbol{\alpha}(k))}$ from Section 7.1 created from $r^{(\emptyset)}$ using the skipping regime. Furthermore we define non-negative function*

$$A : \bigcup_{k\in K\setminus\emptyset} \left(\{k\} \times \mathbb{N}^k\right) \cup \{\emptyset\} \longrightarrow \mathbb{R}_0^+,$$

$$B : \bigcup_{k\in K\setminus\emptyset} \left(\{k\} \times \mathbb{N}^k\right) \cup \{\emptyset\} \longrightarrow \mathbb{R}_0^+.$$

For $k = \emptyset$ we define

$$A(\emptyset) := B(\emptyset) := 1.$$

The generators $V(\mathbf{n})$ describe break down rates

$$v(\mathbf{n}; k, m) := \frac{A(m, n_j : j \in m)}{A(k, n_j : j \in k)}, \qquad k \subsetneq m,$$

and repair rates

$$v(\mathbf{n}; k, m) := \frac{B(k, n_j : j \in k)}{B(m, n_j : j \in m)}, \qquad k \supsetneq m.$$

Here we use $\frac{0}{0} = 0$.

We assume that the customers cannot change environment when they leave the queueing system:

$$R_j(\mathbf{n}) = I \qquad \forall j \in \overline{J}, \mathbf{n} \in \mathbb{N}^{\overline{J}}.$$

Further we define $\beta(k) := 1$ for all $k \in K$ and

$$\gamma_j(k) = \begin{cases} 0 & \text{if} \quad j \in k \\ 1 & \text{if} \quad j \notin k \end{cases} \qquad \forall k \in K.$$

Therefore

$$\alpha_j(k) = \gamma_j(k) = \begin{cases} 0 & \text{if} \quad j \in k \\ 1 & \text{if} \quad j \notin k \end{cases} \qquad \forall k \in K.$$

Then the steady-state distribution π has the form

$$\pi(\mathbf{n}, k) = C^{-1} \prod_{j=1}^{J} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} \frac{A(k, n_j : j \in k)}{B(k, n_j : j \in k)} \tag{8.0.25}$$

with normalization constant

$$C^{-1} := \sum_{\mathbf{n}\in\mathbb{N}_0^{\overline{J}}} \prod_{j=1}^{J} \prod_{\ell=1}^{n_j} \frac{\eta_j}{\mu_j(\ell)} \left(\sum_{k\in K} \frac{A(k, n_j : j \in k)}{B(k, n_j : j \in k)} \right) < \infty.$$

Proof. Let $\boldsymbol{\eta}$ is a unique positive invariant measure of the equation $\boldsymbol{\eta} r^{(\emptyset)} = \boldsymbol{\eta}$. Due to the construction of the matrices $r^{(k)}$ from $r^{(\emptyset)}$ using skipping rule $(\alpha_j(k)\eta_j : j \in \overline{J}_0)$ is an invariant measure of $r^{(k)}$. We define

$$\theta(\mathbf{n}, k) := \frac{A(k, n_j : j \in k)}{B(k, n_j : j \in k)}$$

and show that it solves the equation $\theta(\mathbf{n})Q_{\text{red}}(\mathbf{n}) = 0$. Instead of the matrix representation of $Q_{\text{red}}(\mathbf{n})$ in (8.0.7) we show that $\theta(\mathbf{n},k)$ solves the equations (8.0.14) for all $\mathbf{n}$ and k

$$\begin{aligned}
&0 = \sum_{m\in K} \theta(\mathbf{n};m)\Big(v(\mathbf{n};m,k) + \sum_{j\in\overline{J}} \eta_j\gamma_j(m) r^{(\boldsymbol{\alpha}(m))}(j,0)(\underbrace{R_j(\mathbf{n}+\mathbf{e}_j;m,k)}_{=\delta_{mk}} - \delta_{mk})\Big) \\
\iff &0 = \sum_{m\in K} \theta(\mathbf{n};m) v(\mathbf{n};m,k). \qquad (8.0.26)
\end{aligned}$$

The last equation (8.0.26) holds because for any $k \subsetneq m$

$$\begin{aligned}
\theta(\mathbf{n};k)v(\mathbf{n};k,m) &= \frac{\cancel{A(k,n_j : j\in k)}}{B(k,n_j : j\in k)} \cdot \frac{A(m,n_j : j\in m)}{\cancel{A(k,n_j : j\in k)}} = \frac{A(m,n_j : j\in m)}{B(k,n_j : j\in k)} \cdot 1 \\
&= \frac{A(m,n_j : j\in m)}{B(k,n_j : j\in k)} \cdot \frac{B(k,n_j : j\in m)}{B(k,n_j : j\in m)} = \theta(\mathbf{n},m)v(\mathbf{n};m,k).
\end{aligned}$$

Finally for any $\theta(\mathbf{n},k)$ it holds

$$\theta(\mathbf{n},k) := \frac{A(k,n_j : j\in k)}{B(k,n_j : j\in k)} = \theta(\mathbf{n}+e_j,k), \qquad \forall j \notin k.$$

Because $k = B(\gamma(k))$, vector $\theta(\mathbf{n})$ satisfies the condition (8.0.8).

The vector θ is not zero: By definition of function $A(\emptyset)$ and $B(\emptyset)$ the value $\theta(\mathbf{n},\emptyset)$ is 1.

According to Theorem 8.0.9 for the steady-state solution of the system considered in this corollary it holds $\pi(\mathbf{n},k) = C^{-1}\xi(n)\theta(k)$ with normalization constant C^{-1}. □

Corollary 8.0.20 ([SD03a, Theorem 5.2]). *Consider an ergodic degradable Jackson-Network on a node set $\overline{J}$ satisfying the condition of Theorem 8.0.9. The environment state space K, generators $V(\mathbf{n})$, stochastic matrices $R_j(\mathbf{n})$, functions $\boldsymbol{\gamma}(k)$ and $\beta(k)$ as in Corollary 8.0.19, irreducible reversible routing matrix $r^{(\emptyset)}$. The rerouting matrix $r^{(k)}$ is $r^{(\boldsymbol{\alpha}(k))}$ from Section 7.2, constructed by applying the randomize reflection regime to $r^{(\emptyset)}$. Then the steady-state distribution of the system $\pi(n,k)$ is* (8.0.25) *from Corollary 8.0.19.*

Proof. The proof is the same as in Corollary 8.0.19, with the only difference, that we have a reversible matrix $r^{(\emptyset)}$ and construct $r^{(k)}$ using randomized reflection. We define again η to be a unique solution of the steady-state equation $\boldsymbol{\eta}r^{(\emptyset)} = \boldsymbol{\eta}$ with $\eta_0 = \lambda$. Due to construction of rerouting matrices $r^{(k)}$, for each k it has an invariant measure $(\alpha_j(k)\eta_j : j \in \overline{J}_0)$. The rest of the proof is the same as in Corollary 8.0.19. □

A. Inversion of $\lambda I_W - V$

A.1. General results

In our main Theorem 3.1.16 in part (c) we required the matrix $(\lambda I_W - V)$ to be invertible. $A.2.2$ and $A.2.3$ provided the most important general framework for the invertibility, referring to technical lemmata for matrix algebra which will be proved now.

These technical lemmata do not require the respective matrices to be irreducible. Since we are interested in systems with reducible matrices V which appear e.g. in inventory models (see Example 3.1.18, Proposition 3.3.1 and Proposition 3.3.2), we have to modify especially the proof of invertibility for finite irreducible matrices which can be found e.g. in [Kan05, Lemma 4.12]. The first lemma is the key to the invertibility property in case of finite K, and is therefore of independent interest.

The following definition is equivalent to the definition of essentially diagonally dominant matrix from [Hac91, Hac94, Definition 6.4.8] for finite matrices. We adapted it to our notations and problems.

Definition A.1.1. Let M be a matrix from $\mathbb{K}^{K\times K}$ where the set K of indices is partitioned according to $K = K_W \uplus K_B$, $K_W \neq \emptyset$, whose diagonal elements have following properties:

$$|M_{kk}| = \sum_{m\in K\setminus\{k\}} |M_{km}| \qquad \forall k \in K_B, \tag{A.1.1}$$

$$|M_{kk}| > \sum_{m\in K\setminus\{k\}} |M_{km}| \qquad \forall k \in K_W, \tag{A.1.2}$$

and the <u>flow condition</u> holds

$$\forall \tilde{K}_B \subset K_B,\ \tilde{K}_B \neq \emptyset : \quad \exists \quad k \in \tilde{K}_B, \quad m \in \tilde{K}_B^c : \quad M_{km} \neq 0. \tag{A.1.3}$$

Then we call the matrix M *essentially diagonally dominant*.

Remark A.1.2. Consider the directed transition graph of M, with vertices K and edges $\mathcal{E}$ defined by $km \in \mathcal{E} \iff K_{km} > 0$. Then the condition (A.1.3) guarantees the existence of a path from any vertex in K_B to a some vertex in K_W.

Lemma A.1.3. *Let $M \in \mathbb{K}^{K\times K}$ be an essentially diagonal dominant matrix with $|K| < \infty$ then M is invertible.*

Proof. We prove the lemma by contradiction, and let $x = (x_k : k \in K)$ be a vector with

$$Mx = 0 \text{ with } x \neq 0. \tag{A.1.4}$$

The property $Mx = 0$ leads for all $k \in K$ to

$$\begin{aligned}
-M_{kk}x_k &= \sum_{m\in K\setminus\{k\}} M_{km}x_m \\
\Longrightarrow |M_{kk}||x_k| &\leq \sum_{m\in K\setminus\{k\}} |M_{km}||x_m| \\
\Longrightarrow |M_{kk}|\frac{|x_k|}{||x||_\infty} &\leq \sum_{m\in K\setminus\{k\}} |M_{km}| \underbrace{\frac{|x_m|}{||x||_\infty}}_{\leq 1} \leq \sum_{m\in K\setminus\{k\}} |M_{km}|.
\end{aligned} \tag{A.1.5}$$

We denote by J the set of indices of elements x_k of x with the largest absolute value

$$J := \{k \in K|\ |x_k| = ||x||_\infty\}.$$

Because of $x \neq 0$ and $|K| < \infty$ the set J is non empty.

First we show that

$$\forall k \in K_W : |x_k| < ||x||_\infty \tag{A.1.6}$$

holds, which implies

$$K_W \subset J^c. \tag{A.1.7}$$

For $K_B = \emptyset$ the proof is complete because we have

$$K = K_W \subseteq J^c \subsetneqq K,$$

and so we proceed with the proof for $K_B \neq \emptyset$.

From (A.1.5) and (A.1.2) it follows for all $k \in K_W$

$$\begin{aligned}
|M_{kk}|\frac{|x_k|}{||x||_\infty} &\leq \sum_{m\in K\setminus\{k\}} |M_{km}| < |M_{kk}| \\
\Longrightarrow |M_{kk}|\frac{|x_k|}{||x||_\infty} &< |M_{kk}|.
\end{aligned} \tag{A.1.8}$$

The inequality (A.1.8) is valid if and only if $\frac{|x_k|}{||x||_\infty}$ is strictly less than 1, which implies $|x_k| < ||x||_\infty$ and therefore (A.1.7).

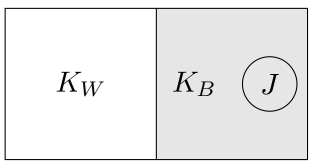


Figure A.1.1.: Sets in Lemma A.1.3. The set K_B is gray.

Next, we analyze the set $J \subset K_B$. For $k \in J$ we examine the kth row of the equation $Mx = 0$.

For all $k \in J$ it follows from (A.1.5)

$$|M_{kk}| \leq \sum_{m \in K\setminus\{k\}} |M_{km}| \underbrace{\frac{|x_m|}{||x||_\infty}}_{\leq 1} \leq \sum_{m \in K\setminus\{k\}} |M_{km}| \leq |M_{kk}|$$

$$\implies \sum_{m \in K\setminus\{k\}} |M_{km}| \underbrace{\frac{|x_m|}{||x||_\infty}}_{\leq 1} = \sum_{m \in K\setminus\{k\}} |M_{km}|. \tag{A.1.9}$$

Because $\frac{|x_m|}{||x||_\infty}$ is strictly less than 1 for all $m \in J^c$, the inequality (A.1.9) yields

$$M_{km} = 0, \qquad \forall k \in J, m \in J^c.$$

Since $K_W \subset J^c$ we have a contradiction to the existence of a path of positive values M_{km} from $k \in J \subset K_B$ to K_W which is guaranteed by (A.1.3). $\square$

Example A.1.4. This example provides a matrix M which fulfills the requirements of Lemma A.1.3 and is therefore invertible It is neither irreducible nor strictly diagonal dominant. We set $\lambda > 0$, $\nu > 0$, all other entries are zero. Figure A.1.2 on page 198 shows the resulting flow graph according to Remark A.1.2.

$M =$

$$\left(\begin{array}{c|cccccc} & 1 \in K_W & 2 \in K_W & 3 \in K_B & 4 \in K_B & 5 \in K_B & 6 \in K_B \\ \hline 1 \in K_W & \lambda & & & & & \\ 2 \in K_W & & (\lambda+\nu) & -\nu & & & \\ 3 \in K_B & & -\nu & \nu & & & \\ 4 \in K_B & & & -\nu & 2\nu & & -\nu \\ 5 \in K_B & & & & -\nu & \nu & \\ 6 \in K_B & & & -\nu & & & \nu \end{array}\right).$$

Note, that this matrix is of the form $M = \lambda I_W - V$ with $V =$

$$\left(\begin{array}{c|cccccc} & 1 \in K_W & 2 \in K_W & 3 \in K_B & 4 \in K_B & 5 \in K_B & 6 \in K_B \\ \hline 1 \in K_W & 0 & & & & & \\ 2 \in K_W & & -\nu & \nu & & & \\ 3 \in K_B & & \nu & -\nu & & & \\ 4 \in K_B & & & \nu & -2\nu & & \nu \\ 5 \in K_B & & & & \nu & -\nu & \\ 6 \in K_B & & & \nu & & & -\nu \end{array}\right),$$

and fits therefore exactly into the realm of our investigations of loss systems in a random environment.

For infinite K we have the following results.

Proposition A.1.5. *Let $M \in \mathbb{R}^{K\times K}$, be a linear operator on $\ell_\infty(\mathbb{R}^K)$. If for all $k \in K$ holds $|M_{kk}| \geq \sum_{m\in K\setminus\{k\}} |M_{km}| + \varepsilon$ for some $\varepsilon > 0$ and $\sup_{k\in K} |M_{kk}| < \infty$,then M is invertible.*

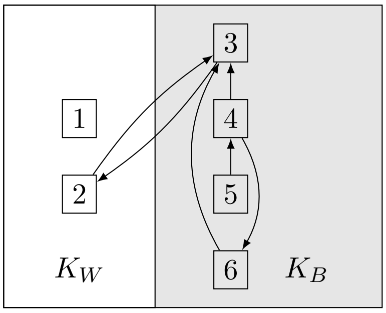


Figure A.1.2.: Graph from example according to Remark A.1.2.

Proof. **(1)** Assume $M_{kk} > 0$ for all $k \in K$. Define $\beta := \frac{1}{\sup_{k\in K} M_{kk}}$, then it holds

$$||I - \beta M||_\infty = \sup_{k\in K} \left(|1 - \underbrace{\beta M_{kk}}_{\leq 1}| + \beta \underbrace{\sum_{m\in K\setminus\{k\}} |M_{km}|}_{\leq M_{kk}-\varepsilon} \right) \tag{A.1.10}$$

$$\leq \sup_{k\in K} (1 - \beta M_{kk} + \beta(M_{kk} - \varepsilon)) < 1. \tag{A.1.11}$$

Thus M is invertible and it holds

$$M^{-1} = \beta \sum_{n=0}^{\infty} (I - \beta M)^n.$$

(2) We define a matrix S with

$$S_{km} = \begin{cases} 1 & \text{for } k = m,\ M_{kk} > 0, \\ -1 & \text{for } k = m,\ M_{kk} < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{A.1.12}$$

Then S is a bounded invertible operator with $S^{-1} = S$. According to **(1)** SM is invertible and it holds $M^{-1} = (SSM)^{-1} = (SM)^{-1}S^{-1} = (SM)^{-1}S$. $\square$

Lemma A.1.6. *Let $M \in \mathbb{R}^{K\times K}$, be a linear operator on $\ell_\infty(\mathbb{R}^K)$ where the set of indices is partitioned according to $K = K_W \uplus K_B$, $K_W \neq \emptyset$, and $|K_B| < \infty$, with the following properties:*

Flow condition*: Define a directed graph $(K, \mathcal{E})$ by*

$$(k, m) \in \mathcal{E} :\Leftrightarrow M(k, m) \neq 0.$$

Then for any $k \in K_B$ there exists some $m = m(k) \in K_W$ such that there exists a directed path of finite length in $(K, \mathcal{E})$ from k to m.

$$\textit{The sequence} \quad |M_{mm}|, m \in K, \quad \textit{is bounded.} \tag{A.1.13}$$

$$|M_{kk}| = \sum_{m \in K \setminus \{k\}} |M_{km}|, \qquad \forall k \in K_B. \tag{A.1.14}$$

$$\sup_{k \in K_W} \sum_{m \in K \setminus \{k\}} |M_{km}| =: ND(K_W) < \infty. \tag{A.1.15}$$

There exists some $\varepsilon(K_W) > 0$ such that

$$\inf_{m \in K_W} |M_{mm}| = ND(K_W) + \varepsilon(K_W) \tag{A.1.16}$$

holds.

Then M is injective.

Proof. In the case $K_B = \emptyset$ the matrix M is strictly diagonal dominant and thus invertible according to Proposition A.1.5.

Let $x = (x_k : k \in K) \in \ell_\infty(\mathbb{R}^K)$ be any vector with

$$Mx = 0 \text{ with } x \neq 0. \tag{A.1.17}$$

(a) To show that

$$\forall k \in K_W : |x_k| < ||x||_\infty \tag{A.1.18}$$

holds, is a word-by-word analogue of that property in the proof of Lemma A.1.3.

(b) We show: $\{|x_k| : k \in K_W\}$ is uniformly bounded away from $||x||_\infty$ from below.

The property $Mx = 0$ leads for all $k \in K$ to

$$\begin{aligned} -M_{kk}x_k &= \sum_{m \in K \setminus \{k\}} M_{km}x_m \qquad \Longrightarrow \\ |M_{kk}||x_k| &\leq \sum_{m \in K \setminus \{k\}} |M_{km}||x_m| \leq ||x||_\infty \sum_{m \in K \setminus \{k\}} |M_{km}| \leq ||x||_\infty ND(K_W), \end{aligned}$$

and therefore

$$|x_k| \inf_{m \in K_W} |M_{mm}| \leq ||x||_\infty ND(K_W) \qquad \Longrightarrow$$

$$|x_k| \leq \frac{ND(K_W)}{\inf_{m \in K_W} |M_{mm}|} ||x||_\infty = \left(1 - \underbrace{\frac{\varepsilon(K_W)}{ND(K_W) + \varepsilon(K_W)}}_{\in(0,1)} \right) ||x||_\infty.$$

(c) We show: $J := \{k \in K : |x_k| = ||x||_\infty\} \neq \emptyset$ and $K_W \subset J^c$.

The second property follows from **(b)**, while the first property holds, because the set $\{|x_k| : k \in K_W\}$ is uniformly bounded away from $||x||_\infty$ from below and K_B is finite, so there must exist some $k(0) \in K_B$ where $|x_{k(0)}| = ||x||_\infty$ is attained.

(d) To show that

$$M_{km} = 0, \qquad \forall k \in J, m \in J^c$$

holds, is a word-by-word analogue of that property in the proof of Lemma A.1.3. Therefore the <u>flow condition</u> is violated and we have proved the theorem. □

A.2. Application to $\lambda I_W - V$

The proof of the following lemma and proposition exploit the the structure of matrix $(\lambda I_W - V)$, where λ, I_W, and V are parameters of an ergodic in continuous time loss system . We emphasize that the generator V is not necessarily irreducible.

Lemma A.2.1. *Let Z be ergodic. Then for any non-empty subset $\tilde{K}_B \subset K_B$ the overall V-transition rate from $\tilde{K_B}$ to its complement $\tilde{K}_B^c = K \setminus \tilde{K_B}$ is positive, i.e.,*

$$\forall \quad \tilde{K}_B \subset K_B, \tilde{K}_B \neq \emptyset : \qquad \exists \quad k \in \tilde{K}_B, m \in \tilde{K}_B^c : v(k,m) > 0. \tag{A.2.1}$$

Remark. Consider the directed transition graph of V, with vertices K and edges $\mathcal{E}$ defined by $km \in \mathcal{E} \iff v(k,m) > 0$. Then the condition (A.2.1) guarantees the existence of a path from any vertex in K_B to a some vertex in K_W. See Remark A.1.2.

Proof. (of Lemma A.2.1) Fix $\tilde{K}_B$ and suppose the system is ergodic and it is started with $Z(0) = (0,k)$, for some $k \in \tilde{K}_B$, i.e., with an empty queue and in an environment state k which blocks the arrival process. From ergodicity it follows that for some $m \in K_W$ must hold

$$P(Z(\sigma_1) = (1,m) | Z(0) = (0,k)) > 0,$$

because there is a positive probability for the first arrival of some customer admitted into the system.

Because no arrival is possible if $m \in K_B$, necessarily $m \in K_W$ holds, and because up to σ_1- no departure or arrival could happen, the only possibility to enter m is by a sequence of transitions triggered by V. Because Z is regular this sequence is finite with probability 1. The path from $k \in \tilde{K}_B$ to $m \in K_W$ of the directed transition graph of V contains an edge $k_1k_2 \in \mathcal{E}$ with $k_1 \in \tilde{K}_B$ and $k_2 \in \tilde{K}_B^c$. $\square$

Proposition A.2.2. *Let Z be ergodic with finite environment space K, and V be the associated generator driving the continuous changes of the environment. Then for any $\lambda > 0$ the matrix $(\lambda I_W - V)$ is invertible.*

Proof. Follows from Lemma A.2.1 and Lemma A.1.3. $\square$

Proposition A.2.3. *Let Z be ergodic with environment space K partitioned according to $K = K_W \uplus K_B$, with $K_W \neq \emptyset$, and with $|K_B| < \infty$, and $\lambda > 0$ such that $\lambda I_W - V$ is surjective on $\ell_\infty(\mathbb{R}^K)$.*

Let the generator matrix $V := (v(k,m) : k,m \in K) \in \mathbb{R}^{K \times K}$ be uniformizable, i.e. it holds $\inf_{k \in K} v(k,k) > -\infty$.

Then the matrix $\lambda I_W - V$ is invertible.

Proof. It is immediate, that $\lambda I_W - V$ fulfills the assumptions (A.1.14), (A.1.15), and (A.1.16) of Lemma A.1.6 with $\varepsilon(K_W) = \lambda$. The <u>flow condition</u> holds in this setting from the ergodicity of the continuous time process with arguments similar to those in the proof of Lemma 3.1.1. We conclude that M is injective. $\square$

Bibliography

[AG07] S. Asmussen and P. W. Glynn. *Stochastic Simulation - Algorithms and Analysis*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007. 148

[AN63] B. Avi-Itzhak and P. Naor. Some queueing problems with service station subject to breakdown. *Operations Research*, 11:303–320, 1963. 34

[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. 3

[BDO01] S. Balsamo, V. De Nitto Persone, and R. Onvural. *Analysis of Queueing Networks with Blocking*. Kluwer Academic Publisher, Norwell, 2001. 56

[BK99] O. Berman and E. Kim. Stochastic models for inventory management at service facilities. *Comm. Statist.– Stochastic Models*, 15(4):695–718, 1999. 4

[BM13] S. Balsamo and A Marin. Separable solutions for Markov processes in random environments. *European Journal of Operational Research*, 229(2):391 – 403, 2013. 4, 145, 147, 148

[BM14] j. W. Baek and S. K. Moon. The M/M/1 queue with a production-inventory system and lost sales. *Applied Mathematics and Computation*, 233:534–544, mar 2014. 46

[Bou94] R. J. Boucherie. A characterization of independence for competing Markov chains with applications to stochastic petri nets. *IEEE Trans. Softw. Eng.*, 20(7):536–544, July 1994. 74

[Bre99] P. Bremaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, volume 31 of *Texts in Applied Mathematics*. Springer–Verlag, New York – Heidelberg – Berlin, 1999. 147, 158

[BS83] C. E. Bell and S. Stidham Jr. Individual versus social optimization in the allocation of customers to alternative servers. *Management Science*, 29:831–839, 1983. 164

[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. `doi:10.1080/15326349508807357`. 30, 31

[BvD10] R.J. Boucherie and N. M. van Dijk. *Queueing networks - A fundamental approach*, volume 154 of *International Series in Operations Research & Management Science*. Springer, New York, 2010. 146

[CG06] C.-F. Chiasserini and M. Garetto. An analytical model for wireless sensor networks with sleeping nodes. *IEEE Transactions on Mobile Computing*, 5(12):1706–1718, 2006. 67

[CLWY07] Z. Chen, C. Lin, H. Wen, and H. Yin. An analytical model for evaluating IEEE 802.15.4 CSMA/CA protocol in low-rate wireless application. In *Advanced Information Networking and Applications Workshops, 2007, AINAW '07. 21st International Conference on*, volume 2, pages 899 –904, may 2007. doi:10.1109/AINAW.2007.77. 67

[CM96] R. Chakka and I. Mitrani. Approximate solutions for open networks with breakdowns and repairs. In Kelly. F.P., Zachary. S., 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. 34

[Cog80] R. Cogburn. Markov chains in random environments: The case of Markovian environments. *Annals of Probability*, 8(5):908–916, 1980. 145

[CT81] R. Cogburn and W. C. Torrez. Markov chains in random environments: The case of Markovian environments. *Journal of Applied Probability*, 18:19–30, 1981. 145

[Dad01] 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. 3

[DG92] Y. Dallery and S. B. Gershwin. Manufacturing flow line systems: a review of models and analytical results. *Queueing Systems*, 12:3–94, 1992. 56

[Dij88] N. M. van Dijk. On Jackson's product form with 'jump-over' blocking. *Operations Research Letters*, 7(5):233–235, 1988. 146, 151

[Dij93] N. M. van Dijk. *Queueing Networks and Product Forms – A Systems Approach.* Wiley, Chichester, 1993. 56, 146

[Dij11] 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. 56

[Dos90] B.T. Doshi. Single server queues with vacations. In H. Takagi, editor, *Stochastic Analysis of Computer and Communication Systems*, pages 217 – 267, Amsterdam, 1990. IFIP, North–Holland. 123

[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. 151

[Dsh97] J. Dshalalow. Queueing systems with state dependent paramenters. In Jewgeni Dshalalow, editor, *Frontiers in Queueing: Models and Applications in Science and Engineering*, chapter 4, pages 61–116. CRC Press, Boca Raton, 1997. 109, 110

[Eco03] A. Economou. A characterization of product-form stationary distributions for queueing systems in random environment. In *Proceedings 17th European Simulation Multiconference (United Kingdom, June 2003)*, 2003. 4, 145

[Eco04] A. Economou. Stationary distributions of discrete-time Markov chains in random environment: Exact computations and bounds. *Stochastic Models*, 20(1):103–127, 2004. `doi:10.1081/STM-120028393`. 4, 145

[Eco05] A. Economou. Generalized product-form stationary distributions for Markov chains in random environments with queueing applications. *Advances in Applied Probability*, 37(1):pp. 185–211, 2005. 4, 145, 147, 148, 149, 177, 178

[Eco14] Antonis Economou. Personal communication to Hans Daduna, 2014. 178

[EF98] A. Economou and D. Fakinos. Product form stationary distributions for queueing networks with blocking and rerouting. *Queueing Systems and Their Applications*, 30:251 – 260, 1998. 145, 151, 156

[Fal96] G. Falin. A heterogeneous blocking system in a random environment. *Journal of Applied Probability*, 33:211 – 216, 1996. 7, 145

[GH74] D. Gross and C. M. Harris. *Fundamentals of Queueing Theory.* John Wiley & Sons, Inc., 1974. 111

[GK90] R.J. Gibbens and F.P. Kelly. Dynamic routing in fully connected networks. *IMA Journal of Mathematical Control & Information*, 77-111, 1990. 146

[GKK95] R.J. Gibbens, F.P. Kelly, and P.B. Key. Dynamic alternative routing. In M.E. Steenstrup, editor, *Routing in Communications Networks*, pages 13–47. Prentice Hall, 1995. 146

[GN67] W.J. Gordon and G.F. Newell. Closed queueing networks with exponential servers. *Operations Research*, 15:254–265, 1967. 3

[Hac91] W. Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme.* Taubner Studienbücher. B.G. Teubner, Stuttgart, 1991. 195

[Hac94] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations.* Applied Mathematical Sciences. Springer-Verlag New York, 1994. `doi:10.1007/978-1-4612-4288-8`. 195

[HJV05] P. Haccou, P. Jagers, and V. Vatutin. *Branching Processes: Variation, Growth, and Extinction of Populations.* Cambridge Studies in Adaptive Dynamics. Cambridge University Press, Cambridge, 2005. 6

[HMRT01] B.R. Haverkort, R. Marie, G. Rubino, and K. Trivedi. *Performability Modeling, Technique and Tools.* Wiley, New York, 2001. 148

[HW84] W. Helm and K.-H. Waldmann. Optimal control of arrivals to multiserver queues in a random environment. *Journal of Applied Probability*, 21:602–615, 1984. 7

[Jac57] J.R. Jackson. Networks of waiting lines. *Operations Research*, 5:518–521, 1957. viii, 3, 73, 145, 163, 215, 217

[JW10] L. Jiang and J. Walrand. A distributed csma algorithm for throughput and utility maximization in wireless networks. *Networking, IEEE/ACM Transactions on*, 18(3):960 –972, june 2010. doi:10.1109/TNET.2009.2035046. 67

[Kan05] C. Kanzow. *Numerik linearer Gleichungssysteme: Direkte und iterative Verfahren*. Springer, 2005. 195

[KD12] R. Krenzler and H. Daduna. Loss systems in a random environment - steady state analysis. http://preprint.math.uni-hamburg.de/public/papers/prst/prst2012-04.pdf, November 2012. 68, 93, 219

[KD13a] R. Krenzler and H. Daduna. Loss systems in a random environment. http://arxiv.org/abs/1312.0539, December 2013. 219

[KD13b] R. Krenzler and H. Daduna. Loss systems in a random environment - embedded markov chains analysis. http://preprint.math.uni-hamburg.de/public/papers/prst/prst2013-02.pdf, May 2013. 68, 92, 94, 219

[KD14] R. Krenzler and H. Daduna. Modeling and performance analysis of a node in fault tolerant wireless sensor networks. In K. Fischbach and U.R. Krieger, editors, *Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault-Tolerance*, pages 73–78, Heidelberg, 2014. GI/ITG, Springer. 66, 145, 219

[KD15a] R. Krenzler and H. Daduna. Loss systems in a random environment: steady state analysis. *Queueing Systems*, 80(1-2):127–153, 2015. URL: http://dx.doi.org/10.1007/s11134-014-9426-6, doi:10.1007/s11134-014-9426-6. 9, 17, 19, 23, 26, 46, 48, 52, 54, 55, 219

[KD15b] R. Krenzler and H. Daduna. Performability analysis of an unreliable M/M/1-type queue. In *Leistungs-, Zuverlässigkeits- und Verlässlichkeitsbewertung von Kommunikationsnetzen und verteilten Systemen : 8. GI/ITG-Workshop MMBnet 2015, 10./11. September 2015*, Berichte des Fachbereichs Informatik der Universität Hamburg, pages 90–95. Universität Hamburg, 2015. URL: http://edoc.sub.uni-hamburg.de/informatik/volltexte/2015/216/. 20, 58, 60, 219

[KDO14] R. Krenzler, H. Daduna, and S. Otten. Randomization for markov vhains with applications to networks in a random environment. http://arxiv.org/abs/1407.8378, July 2014. 219

[KDO16] R. Krenzler, H. Daduna, and S. Otten. Jackson networks in non-autonomous random environments. *Advances in Applied Probability*, 48(2), 2016. to appear in June 2016. 219

[Kei79] J. Keilson. *Markov chain models – Rarity and exponentiality.* Springer, New York, 1979. 26

[Kel76] F. Kelly. Networks of queues. *Advances in Applied Probability*, 8:416–432, 1976. 3

[Kel79] F. P. Kelly. *Reversibility and Stochastic Networks.* John Wiley and Sons, Chichester – New York – Brisbane – Toronto, 1979. 26

[Kes80] H. Kesten. Random processes in random environments. In W. Jäger, H. Rost, and P. Tautu, editors, *Biological Growth and Spread*, volume 38 of *Lecture Notes in Biomathematics*, chapter Ib, pages 82–92. Springer, Berlin, 1980. Proceedings of the Conference on Models of Biological Growth and Spread, University of Heidelberg, 1979. 7

[Kle75] L. Kleinrock. *Queueing Theory*, volume I. John Wiley and Sons, New York, 1975. 111

[Kle76] L. Kleinrock. *Queueing Theory*, volume II. John Wiley and Sons, New York, 1976. 156

[KLM11] A. Krishnamoorthy, B. Lakshmy, and R. Manikandan. A survey on inventory models with positive service time. *OPSEARCH*, 48(2):153–169, 2011. 6, 68, 89, 133

[KMD15] A. Krishnamoorthy, R. Manikandan, and Shajin Dhanya. Analysis of a multiserver queueing-inventory system. *Advances in Operations Research*, 2015, 2015. `doi:10.1155/2015/747328`. 46

[KML13] A. Krishnamoorthy, R. Manikandan, and B. Lakshmy. A revisit to queueing-inventory system with positive service time. *Annals of Operations Research*, 2013. 21, 46, 120

[KN13] A. Krishnamoorthy and Viswanath C. Narayanan. Stochastic decomposition in production inventory with service time. *European Journal of Operational Research*, 228:358–366, July 2013. 21, 46, 145

[KPC12] A. Krishnamoorthy, P.K. Pramod, and S.R. Chakravarthy. Queues with interruptions: a survey. *TOP*, 2012. `doi:10.1007/s11750-012-0256-6`. 145

[KY12] V. Kulkarni and K. Yan. Production-inventory systems in stochastic environment and stochastic lead times. *Queueing Systems and Their Applications*, 70:207–231, 2012. 145

[Li11] W.W. Li. Several characteristics of active/sleep model in wireless sensor networks. In *New Technologies, Mobility and Security (NTMS), 2011 4th IFIP International Conference on New Technologies, Mobility and Security*, pages 1–5, feb. 2011. `doi:10.1109/NTMS.2011.5721073`. 66, 67

[Lig85] T.M. Liggett. *Interacting Particle Systems*, volume 276 of *Grundlehren der mathematischen Wissenschaften.* Springer, Berlin, 1985. 156

[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. 4, 27, 28, 29, 38

[LTL05] J. Liu and T. Tong Lee. A framework for performance modeling of wireless sensor networks. In *Communications, 2005. ICC 2005. 2005 IEEE International Conference on Communications*, volume 2, pages 1075 – 1081 Vol. 2, 2005. doi:10.1109/ICC.2005.1494513. 66, 67, 68, 76

[MA68] I. Mitrani and B. Avi-Itzhak. A many-server queue with service interruptions. *Operations Research*, 16:628–638, 1968. 34

[MAG06] M.K. Mehmet Ali and H. Gu. Performance analysis of a wireless sensor network. In *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, volume 2, pages 1166 –1171, april 2006. doi:10.1109/WCNC.2006.1683635. 67

[Mey84] J.F. Meyer. Performability modeling of distributed real-time systems. In G. Iazeolla, P.J. Courtois, and A. Hordijk, editors, *Mathematical Computer Performance and Reliability*, pages 361–372, Amsterdam, 1984. NOrth-Holland. Proceding of the International Workshop, Pisa, Italy, September 26 - 30, 1983. 148

[Neu81] M.F. Neuts. *Matrix Geometric Solutions in Stochastic Models - An Algorithmic Approach.* John Hopkins University Press, Baltimore, MD, 1981. 6, 27, 38, 132

[Neu89] M.F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and Their Applications.* Marcel Dekker, New York, 1989. 6

[NSB+03] A. Nucci, B. Schroeder, S. Bhattacharyya, N. Taft, and C. Diot. IGP link weight assignment for transient link failures. In J. Charzinski, R. Lehnert, and P. Tran-Gia, editors, *Providing Quality of Service in Heterogeneous Environments*, volume 5a of *Teletraffic Science and Engineering*, pages 321–330, Amsterdam, 2003. Proceedings of the 18th International Teletraffic Congress, Elsevier. 146

[O'C93] Colm Art O'Cinneide. Entrywise perturbation theory and error analysis for Markov chains. *Numerische Mathematik*, 65(1):109–120, 1993. URL: http://dx.doi.org/10.1007/BF01385743, doi:10.1007/BF01385743. 37

[Onv90] R. O. Onvural. Closed queueing networks with blocking. In H. Takagi, editor, *Stochastic Analysis of Computer and Communication Systems*, pages 499–528. North-Holland, Amsterdam, 1990. 56

[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. 56, 57, 81

[Pes73] P.H. Peskun. Optimum Monte-Carlo sampling using Markov chains. *Biometrika*, 60:607–612, 1973. 160

[QFX+11] T. Qiu, L. Feng, F. Xia, G. Wu, and Y. Zhou. A packet buffer evaluation method exploiting queueing theory for wireless sensor networks. *Computer Science and Information Systems*, 8(4):1027–1049, 2011. 67

[RT96] V. Ramaswami and P.G. Taylor. An operator analytic approach to product-form networks. *Communications in Statistics. Stochastic Models*, 12(1):121–142, 1996. doi:10.1080/15326349608807376. 4, 30, 31, 33, 34

[SAH13] M. Saffari, S. Asmussen, and R. Haji. The M/M/1 queue with inventory, lost sale and general lead times. *Queueing Systems*, 75:65–77, September 2013. 4, 5, 13, 47, 53, 54, 145

[Sau06] C. Sauer. *Stochastic product form networks with unreliable nodes: Analysis of performance and availability*. PhD thesis, University of Hamburg, Department of Mathematics, 2006. viii, 4, 6, 58, 215, 216, 217, 218

[Sch73] R. Schassberger. *Warteschlangen*. Springer, Wien, 1973. 48

[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. 151

[Sch04] M. Schwarz. *Stochastic Models in Inventory Theory*. Shaker Verlag, Aachen, 2004. viii, 112, 215, 217

[SD03a] C. Sauer and H. Daduna. Availability formulas and performance measures for separable degradable networks. *Economic Quality Control*, 18:165–194, 2003. 4, 5, 6, 54, 55, 67, 145, 146, 147, 151, 156, 190, 191, 192, 194

[SD03b] C. Sauer and H. Daduna. Separable networks with unreliable servers. In J. Charzinski, R. Lehnert, and P. Tran-Gia, editors, *Providing QoS in Heterogeneous Environments*, volume 5b of *Teletraffic Science and Engineering*, pages 821–830. Elsevier Science, Amsterdam, 2003. 67

[SD05] M. Schwarz and H. Daduna. Queueing systems with inventory management under random lead times and with backordering. Preprint 2005-03, Schwerpunkt Mathematische Statistik und Stochastische Prozesse, Fachbereich Mathematik der Universität Hamburg, 2005. 35

[Ser99] R. F. Serfozo. *Introduction to Stochastic Networks*, volume 44 of *Applications of Mathematics*. Springer, New York, 1999. 151

[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. 4, 5, 47, 53

[SS12] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *Annals of Applied Probability*, 22(1):128–171, 2012. 146, 148, 161

[SSD+06] M. Schwarz, C. Sauer, H. Daduna, R. Kulik, and R. Szekli. M/M/1 queueing systems with inventory. *Queueing Systems and Their Applications*, 54:55–78, 2006. 4, 5, 12, 27, 43, 45, 47, 67, 106, 109, 145

[Sti09] S. Stidham Jr. *Optimal design of queueing systems.* CRC Preess, Boca Raton, 2009. 148, 164

[Ter94] R.J. Tersine. *Principles of Inventory and Materials Management.* PTR Prentice Hall, Englewood Cliffs, N.J., 4 edition, 1994. 43

[Tie98] L. Tierney. A note on Metropolis-Hastings kernels for general state spaces. *Annals of Applied Probability*, 8:1–9, 1998. 160

[TOKB02] G.S. Tsitsiashvili, M.A. Osipova, N.V. Koliev, and D. Baum. A product theorem for Markov chains with application to PF-queueing networks. Preprint, University of Trier, Department IV, Trier, Germany, 2002. 4, 145, 146, 147, 149

[Vin08] K. Vineetha. *Analysis of inventory systems with positive and / negligible service time.* PhD thesis, University of Calicut, Department of Statistics, January 2008. 4, 5, 6, 89, 109, 120, 121, 140

[WC58] H. White and L.S. Christie. Queueing with preemptive priorities or with breakdown. *Operations Research*, 6:79–95, 1958. 34

[WDW07] Y. Wang, H. Dang, and H. H. Wu. A survey on analytic studies of Delay-Tolerant Mobile Sensor Networks. *Wireless Communications and Mobile Computing*, 7(10):1197–1208, 2007. 66

[Whi85] P. Whittle. Scheduling and characterization problems for stochastic networks. *Journal of the Royal Statistical Society, Series B*, 47(3):407–415, 1985. 148, 164

[WWDL07] H. Wu, H., Y. Wang, H. Dang, and F. Lin. Analytic, Simulation, and Empirical Evaluation of Delay/Fault-Tolerant Mobile Sensor Networks. *IEEE Transactions on Wireless Communications*, 6(9):3287–3296, 2007. 66, 67, 68, 69, 70, 74, 75, 76

[Yec73] U. Yechiali. A queuing-type birth-and-death process defined on a continuous-time markov chain. *Operations Research*, 21(2):604–609, 1973. 145

[YM95] G. Yamazaki and M. Miyazawa. Decomposability in queues with background states. *Queueing Systems and Their Applications*, 20:453–469, 1995. 145, 148, 149

[Zhu94] Y. Zhu. Markovian queueing networks in a random environment. *OR Letters*, 15:11 – 17, 1994. 4, 145, 146, 147, 148, 177, 178

[ZL11] Y. Zhang and W. Li. An energy-based stochastic model for wireless sensor networks. *Wireless Sensor Networks*, 3(9):322–328, 2011. 66, 67, 68, 76

[Zol66] V. M. Zolotarev. Distribution of queue length and number of operating lines in a system of Erlang type with random breakage and restoration of lines. *Selected Translations in Mathematical Statistics and Probability*, 6:89–99, 1966. 145

[ZW09] J. Zhang and Q. Wang. A queuing network model for shovel-truck-crusher systems in open-pit mining. In *APCOM 2009*, 2009. 79, 85, 86

Index

Addenda required by §7 of the doctoral degree regulations of the MIN Faculty

Abstract

Queueing networks with product-form steady-state distribution have found many fields of applications, e.g. production systems, telecommunications, and computer system modeling. The success of this class of models and its relatives stems from the simple structure of the steady-state distribution which provides access to easy performance evaluation procedures. Starting from the work of Jackson [Jac57] various generalizations have been developed.

In real world queueing systems are not isolated and interact with their environment. Adding a random environment to a model usually makes the model more realistic but also more complex to analyze. Nevertheless, under some conditions it is still possible to obtain analytical results. A branch of research which recently has found interest are queueing networks in a random environment with product form steady-state distributions.

The main theoretical contributions of this thesis are twofold: (i) We develope a general theory that comprise models with stationary product-form distribution in inventory theory in [Sch04] and Jackson networks with unreliable nodes with stationary product-form distribution in [Sau06]. An important property of the resulting general model is that the queueing system and the environment interact in both directions: the queues can influence the environment and the environment can influences the queues. (ii) With respect to applications we show how different models known from literature can be interpreted in terms of the general theory, construct new models in various applications, and develope an approximation method.

In Part I we analyze single-queue systems. In Section 1 we introduce a *loss system*. In Section 2 we generalize product form lost-sales inventory models from [Sch04] and several other published papers with related models as a loss system with exponential service time. The term *loss* means that customers get lost when the environment stays in some special states – the blocking states. In Section 2.1.4 we develop an approximation method for system without loss of customers based on loss systems. In Section 2.2 we apply our loss system results in fields different from inventory management: we analyze in detail an unreliable server with preventive maintenance in Section 2.2.4, a node of a wireless sensor network in Section 2.2.5, and a crusher station in open-pit mining in Section 2.2.6.

In Section 3 we analyze the Markov chain embedded at departure instants of the loss system. The embedded Markov chains are an important tool for analyzing queueing system with general service times – the $M/G/1/\infty$ queues. The famous and frequently used result in classical $M/G/1/\infty$ theory is that the steady-state distribution of an $M/G/1/\infty$ system as continuous time process and as embedded Markov chain, observed at departure times, are the same. We show that this is in general not true for the steady-state

distribution of loss systems. We use an embedded Markov chain analysis to extend our results from Section 2 to some loss systems with general service times.

In Part II we extend our results for a single-queue loss system to Jackson networks in a random environment. We replace the concept of loss of customers by special rerouting regimes. We establish a connection between these rerouting regimes and randomized random walks. In Section 8 we consider systems where the interaction between environment and queuing system depend on the number of customers in the system. This extension finally allows us to include results about Jackson networks with unreliable nodes from [Sau06] as special cases.

Zusammenfassung

Warteschlangennetze, deren stationäre Verteilung eine Produktform hat, sind in unterschiedlichen Bereichen angewendet worden, zum Beispiel: Produktionssysteme, Telekommunikation und Modellierung von Rechnersystemen. Ihren Erfolg haben sie ihrer einfachen Struktur der stationärer Verteilung zu verdanken. Sie vereinfacht die Analyse von Leistungskenngrößen der zu modellierenden Systeme. Aufbauend auf der grundlegenden Arbeit von Jackson [Jac57] wurden weitere Verallgemeinerungen entwickelt.

In der realen Welt sind Warteschlangensysteme nie isoliert. Sie befinden sich in einer Umgebung, mit der sie interagieren. Das Hinzufügen einer Umgebung zu einem reinen Warteschlangenmodell führt häufig zu einem besseren Gesamtmodell. Gleichzeitig wird die mathematische Analyse dieses Modells schwieriger und komplexer. Unter speziellen Bedingungen ist es dennoch möglich, für derartige Systeme analytische Lösungen zu erhalten. Viele solche Systeme, gehören zu Warteschlangennetzwerken in einer zufälligen Umgebung, deren stationäre Verteilung eine Produktform hat.

Die Hauptbeiträge dieser Dissertation liegen in zwei Bereichen: (i) Wir entwickeln eine allgemeine Theorie, die Modelle mit stationärer Produktformverteilung aus den Lagerhaltungsmodellen aus [Sch04] und Jackson-Netzwerke mit unzuverlässigen Knoten mit stationärer Produktformverteilung aus [Sau06] gleichzeitig umfaßt. Eine wichtige Eigenschaft dieses allgemeinen Modells ist, dass die Warteschlangensysteme und die Umgebung sich gegenseitig beeinflussen: die Warteschlangen können die Umgebung beeinflussen und die Umgebung die Warteschlangen. (ii) Auf Anwendungen bezogen, zeigen wir, dass viele aus der Literatur bereits bekannte Modelle sich mit Hilfe dieser allgemeinen Theorie darstellen lassen. Wir stellen neue Modelle für unterschiedliche Bereiche vor und entwickeln ein Näherungsverfahren.

In Teil I untersuchen wir Systeme mit einer einzigen Warteschlange. In Abschnitt 1 stellen wir ein *Verlustsystem* (*loss system*) vor . In Abschnitt 2 verallgemeinern wir Lagerhaltungsmodelle mit Kundenverlust aus [Sch04] und mehrere ähnliche publizierte Modelle. Diese Verallgemeinerung bezeichnen wir als Verlustsystem. Der Bergriff *Verlust* (*loss*) bezieht sich auf die Modellannahme, dass Kunden verlorengehen, solange die Umgebung in sogenannten blockierenden Zuständen ist. In Abschnitt 2.1.4 entwickeln wir ein Näherungsverfahren für Systeme ohne Kundenverlust, das auf Systemen mit Kundenverlust basiert. In Abschnitt 2.2 nutzen wir Ergebnisse für Verlustsysteme in anderen Bereichen außerhalb der Lagerhaltung: wir untersuchen detailliert ein Warteschlangensystem mit einem unzuverlässigen Bediener mit präventiver Wartung in Abschnitt 2.2.4, einen Knoten in einem drahtlosen Sensornetzwerk in Abschnitt 2.2.5 und eine Zerkleinerungsanlage im Bergbau in Abschnitt 2.2.6.

In Abschnitt 3 untersuchen wir eingebettete, zu Kundenabgangszeiten beobachtete, Markov-Ketten von Verlustsystemen. Eingebetteten Markov-Ketten sind ein wichtiges Werkzeug für die Untersuchung der Warteschlangensysteme mit allgemeinen Bedienzeitverteilungen, das heißt Warteschlangensysteme vom Typ $M/G/1/\infty$. Ein bekanntes und häufig benutztes Ergebnis in der klassischen $M/G/1/\infty$ Theorie ist, dass für ein $M/G/1/\infty$ System die stationären Verteilungen des Prozesses in stetiger Zeit und dijenige der eingebetteten Markov-Kette, betrachtet zu Kundenabgangszeiten, übereinstimmen. Wir zeigen, dass dies für die stationären Verteilungen eines Verlustsystems im Allgemeinen nicht gilt. Wir benutzen Markov-Ketten-Methoden um unsere Ergebnisse für exponentielle Bedienzeiten aus Abschnitt 2 in einigen Fällen auf Systeme mit allgemeiner

Bedienzeit zu erweitern.

In Teil II erweitern wir unsere Ergebnisse für Verlustsysteme mit einer Warteschlange zu Jackson-Netzwerken in zufälliger Umgebung. Wir ersetzen das Konzept des Kundenverlustes durch spezielle Reroutingregeln. Wir stellen einen Zusammenhang her zwischen unterschiedlichen Reroutingregeln und randomisierten Irrfahrten. Zum Schluss, in Abschnitt 8, erlauben wir zusätzlich, dass die Wechselwirkung zwischen der Umgebung und dem Warteschlangennetz von der Kundenzahl im Gesamtsystem abhängen kann. Diese Erweiterung ermöglicht es, die Ergebnisse über Jackson-Netzwerke mit unzuverlässigen Knoten aus [Sau06] als Spezialfälle der gemeinsamen Theorie zu erfassen.

List of publications derived from the dissertation

Parts of this thesis are or will be published as following preprints and articles:

1. [KD12] – preprint about loss systems in continuous time.
2. [KD13b] – preprint about loss systems analyzed with embedded Markov chains.
3. [KD13a] – preprint. It is extension and unification [KD12] and [KD13b]. We added analysis of unreliable $M/M/1/\infty$ queueing system with control of repair and maintenance.
4. [KD14] – paper for MMB & DFT 2014 conference. It is about application of a loss system for modeling of a node in a wireless network.
5. [KDO14] – preprint. It is an extension of loss-system results for single-queue systems to Jackson networks. We discuss there the connection to random walk problems and sampling methods for simulation of stochastic processes in discrete state space.
6. [KD15a] – journal article, it is a shortened version of [KD12] (systems with constant input rate λ).
7. [KD15b] – paper for MMBnet2015 conference. It contains the *unreliable* $M/M/1/\infty$ *queueing system with control of repair and maintenance* from [KD13a] with some extensions and improvements
8. [KDO16] – journal article. It contains parts from [KDO14] focused on the section *Jackson networks in a random environment.*

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den 09.02.2016 Ruslan Krenzler