

**MARTIN EPP**

**Performance evaluation of shuttle-based  
storage and retrieval systems using discrete-time  
queueing network models**

**BAND 90**

**Wissenschaftliche Berichte des Instituts für Fördertechnik und  
Logistiksysteme des Karlsruher Instituts für Technologie (KIT)**





Martin Epp

Performance evaluation of shuttle-based  
storage and retrieval systems using discrete-time  
queueing network models

WISSENSCHAFTLICHE BERICHTE

Institut für Fördertechnik und Logistiksysteme  
am Karlsruher Institut für Technologie (KIT)

BAND 90

# Performance evaluation of shuttle-based storage and retrieval systems using discrete-time queueing network models

by  
Martin Epp



Dissertation, Karlsruher Institut für Technologie  
KIT-Fakultät für Maschinenbau

Tag der mündlichen Prüfung: 16. Mai 2017

Referenten: Prof. Dr.-Ing. habil. Kai Furmans  
Prof. Dr.-Ing. habil. Thorsten Schmidt

**Impressum**



Karlsruher Institut für Technologie (KIT)  
KIT Scientific Publishing  
Straße am Forum 2  
D-76131 Karlsruhe

KIT Scientific Publishing is a registered trademark  
of Karlsruhe Institute of Technology.  
Reprint using the book cover is not allowed.

[www.ksp.kit.edu](http://www.ksp.kit.edu)



*This document – excluding the cover, pictures and graphs – is licensed  
under a Creative Commons Attribution-Share Alike 4.0 International License  
(CC BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/deed.en>*



*The cover page is licensed under a Creative Commons  
Attribution-No Derivatives 4.0 International License (CC BY-ND 4.0):  
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>*

Print on Demand 2018 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 0171-2772

ISBN 978-3-7315-0742-0

DOI 10.5445/KSP/1000077211





# **Performance evaluation of shuttle-based storage and retrieval systems using discrete-time queueing network models**

Zur Erlangung des akademischen Grades

**Doktor der Ingenieurwissenschaften**

der Fakultät für Maschinenbau  
Karlsruher Institut für Technologie (KIT)  
genehmigte

**Dissertation**

von

**Dipl.-Ing. Martin Epp**

Tag der mündlichen Prüfung:

16.05.2017

Hauptreferent:

Prof. Dr.-Ing. habil. Kai Furmans

Korreferent:

Prof. Dr.-Ing. habil. Thorsten Schmidt



# Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Fördertechnik und Logistiksysteme des Karlsruher Instituts für Technologie. Ich möchte mich an dieser Stelle bei allen Personen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Zunächst möchte ich mich bei Prof. Dr.-Ing. habil. Kai Furmans, Leiter des Instituts für Fördertechnik und Logistiksysteme, für die Übernahme des Haupotreferats bedanken. Seine Vorarbeiten im Bereich der zeitdiskreten Bedientheorie sowie seine Unterstützung meiner Tätigkeiten als wissenschaftlicher Mitarbeiter im Bereich der Leistungsabnahmen von Shuttle-Systemen haben mich zum Erstellen der vorliegenden Arbeit motiviert. Darüber hinaus hat die Möglichkeit zum selbstständigen Arbeiten, gepaart mit zielführenden Diskussionen, zum erfolgreichen Abschluss meiner Dissertation beigetragen.

Prof. Dr.-Ing. habil. Thorsten Schmidt danke ich für die Übernahme des Korreferats sowie Prof. Dr.-Ing. Gisela Lanza für die Übernahme des Prüfungsvorsitzes.

Zudem möchte ich mich bei Prof. Dr. Barış Tan und dem Karlsruher House of Young Scientists (KHYS) bedanken, die meinen dreimonatigen Auslandsaufenthalt an der Koç University in Istanbul gefördert haben. Während dieser Zeit konnte ich mich ausschließlich meiner wissenschaftlichen Arbeit widmen. Die anregenden Diskussionen mit Prof. Tan und seiner Forschungsgruppe sowie die Teilnahme an den Doktorandenseminaren an der Koç University haben sich positiv auf die vorliegende Arbeit ausgewirkt.

Bei meinen aktiven und ehemaligen Kollegen möchte ich mich für die angenehme Arbeitsatmosphäre am Institut und die gegenseitige Unterstützung bedanken. Insbesondere möchte ich mich bei Dr.-Ing. Christian Huber bedanken, der meine Begeisterung für das wissenschaftliche Arbeiten während meiner Zeit als studentische Hilfskraft und Diplomand geweckt hat. Ebenso

danke ich meinen studentischen Hilfskräften und Abschlussarbeitern, im Speziellen Simon Wiedemann, die meine Forschungsarbeit begleitet haben.

Meinen Freunden möchte ich dafür danken, dass ich bei ihnen ein wunderbares soziales Umfeld vorfinden konnte, das auch meine arbeitsintensiven Zeiten ertragen hat. Weiterhin möchte ich mich bei meiner Familie und insbesondere bei meinen Eltern bedanken, die mich auf meinem eingeschlagenen Lebens- und Bildungsweg stets unterstützt und gefördert haben. Mein größter Dank geht an meine Frau Eda für ihre Liebe, ihren Rückhalt und ihre Unterstützung während der Dissertation.

Karlsruhe, November 2017

Martin Epp

# Kurzfassung

Shuttle-Systeme stellen heutzutage einen wichtigen Baustein vieler Lager- und Distributionssysteme dar. Durch die hohen erzielbaren Durchsätze sowie der Möglichkeit zur Anpassung und Erweiterung in Hinblick auf sich ändernde Kundenanforderungen sind sie eine interessante Alternative zu automatischen Kleinteilelagern. Aufgrund der raschen technischen Weiterentwicklung von Shuttle-Systemen sind jedoch nur wenige wissenschaftliche Arbeiten zu der Vielzahl an unterschiedlichen Systemausprägungen vorhanden. Insbesondere fehlen Modelle, welche die Form der Bearbeitungszeitverteilungen berücksichtigen, und mit denen die gesamte Wahrscheinlichkeitsverteilung der Durchlaufzeit von Auslageraufträgen bestimmt werden kann.

Aus diesem Grund wird in dieser Arbeit eine neue Vorgehensweise entwickelt, mit der sich unterschiedliche Ausprägungen von Shuttle-Systemen modellieren lassen. Sie beruht auf der Abbildung der Shuttle-Systeme als zeitdiskrete offene Bediensystemnetzwerke. Mittels eines Dekompositionsansatzes wird das Bediensystemnetzwerk in unabhängige  $G|G|1$  Bediensysteme unterteilt, welche die Ressourcen des Shuttle-Systems darstellen. Damit können die gesamte Durchlaufzeitverteilung und die Verteilung der physisch wartenden Aufträge vor den Ressourcen bestimmt werden.

Dieser Dekompositionsansatz wird im Anschluss an zwei Ausprägungen von Shuttle-Systemen angewendet. Das erste System besteht aus einem Lager mit Fahrzeugen, welche weder Gassen noch Ebenen wechseln können. Das zweite System besteht aus einem Lager mit einem Shuttle je Gasse, welches die Ebenen wechseln kann. In beiden System werden Fahrzeuge modelliert, die mehrere Level mit ihrem Lastaufnahmemittel erreichen können. Ebenso wird die Verbindung der Shuttle-Systeme zu den nachgelagerten Kommissionierstationen abgebildet.

Die Approximationsgüte des Ansatzes wird anhand eines Vergleichs zu einer ereignisdiskreten Simulation überprüft. In den jeweils über 1.000 untersuchten

Ausprägungen der beiden Shuttle-Systeme liefert das Dekompositionsverfahren eine hohe Approximationsgüte. Insbesondere bei Systemausprägungen, die exponentialverteilte Zwischenankunftszeitverteilungen und eine hohe Anzahl Gassen und Ebenen aufweisen, sind die durchschnittlichen Abweichungen sehr gering.

Im Anschluss wird anhand des neu entwickelten Ansatzes gezeigt, wie unter den Systemanforderungen Lagerkapazität, maximale Lagerabmessungen, Durchsatz und 95% Quantil der Durchlaufzeitverteilung eine geeignete Lagerausprägung bestimmt werden kann. Darüber hinaus wird der Einfluss von Fahrzeugen, die mehrere Level erreichen können, auf die Entscheidung der Lagerausprägung dargestellt. Ebenso wird der Einfluss von wiedereintretenden Behältern sowie unterschiedlichen Ausprägungen der Variabilität der vorhandenen Zufallsvariablen auf die Durchlaufzeitverteilung untersucht.

# Abstract

Shuttle-based storage and retrieval systems (SBS/RSs) are an important part of today's warehouses. Due to their ability to generate a high throughput and to adjust to different customer demands, they represent an interesting alternative to traditional automated storage and retrieval systems (AS/RSs) for mini-loads. The fast technological development in the field of SBS/RSs creates the need for performance evaluation tools that can be used during the decision making process. In particular, analytical models are missing that consider as input the shape of the processing time distributions and provide as output the retrieval transaction time distribution.

Hence, in this work we develop a new approach that can be applied to model different configurations of SBS/RSs. The approach is based on the modeling of SBS/RSs as discrete-time open queueing networks. Afterward, the network is decomposed into independent  $G|G|1$  queueing systems, representing the resources of the system. The approach yields the complete probability distributions of the retrieval transaction time and the number of storage transactions waiting in front of the resources.

Subsequently, both a tier-captive and a tier-to-tier SBS/RS are modeled using the approach. In these models, the load handling devices of the shuttles are able to reach one or more than one level of the storage rack, respectively. In addition, we model the succeeding picking process and the re-entrance of the non-empty bins to the SBS/RS.

The approximation quality of the decomposition approach is tested against a discrete-event simulation. We create for both the tier-captive and tier-to-tier SBS/RSs over 1,000 different system configurations, and compare the performance measures of our approach to the values obtained by the simulation. In general, the decomposition approach reaches a high approximation quality. Especially in system configurations with Poisson arrivals and a large number of aisles and tiers, the average deviations are very low.

## Abstract

---

Finally, we demonstrate how to use our approach during the decision making process. Given the system requirements storage capacity, maximum floor space and height, throughput, and service level in terms of the 95% quantile of the retrieval transaction time distribution, we show how to determine a suitable system design that fulfills the requirements. Moreover, we demonstrate the impact of multi-level shuttles, re-entrant bins, and the variability of the random variables on the design decision.

# Contents

<b>Kurzfassung</b> . . . . .	iii
<b>Abstract</b> . . . . .	v
<b>1 Introduction</b> . . . . .	1
1.1 Problem description . . . . .	3
1.2 Organization of the thesis . . . . .	5
<b>2 Characterization of SBS/RSSs</b> . . . . .	9
2.1 System load – flow of storage and retrieval transactions . . . . .	13
2.2 Physical design . . . . .	15
2.3 Control policies . . . . .	22
2.4 Design requirements and performance measures . . . . .	27
<b>3 Literature review</b> . . . . .	31
3.1 Publications on AVS/RSSs . . . . .	31
3.1.1 Analytical models . . . . .	31
3.1.2 Simulation studies . . . . .	37
3.2 Publications on SBS/RSSs . . . . .	39
3.2.1 Analytical models . . . . .	39
3.2.2 Simulation studies . . . . .	43
3.3 Conclusion of the literature review . . . . .	45
<b>4 Decomposition approach</b> . . . . .	47
4.1 Basics of discrete-time probability theory . . . . .	47
4.2 Models and methods in the discrete time domain . . . . .	49
4.2.1 $G G 1$ queueing system . . . . .	55
4.2.2 Stochastic merge . . . . .	57
4.2.3 Stochastic split . . . . .	59

4.2.4	Fast split approximation . . . . .	62
4.3	General modeling approach . . . . .	64
4.3.1	Steps of the modeling approach . . . . .	66
4.3.2	Assumptions . . . . .	70
<b>5</b>	<b>Modeling of tier-captive and tier-to-tier configurations . . . . .</b>	<b>73</b>
5.1	Tier-captive configuration . . . . .	73
5.1.1	System description . . . . .	73
5.1.2	Queueing network model . . . . .	81
5.1.3	Modeling of the service time distributions . . . . .	86
5.1.4	Computation of the network performance measures . . . . .	96
5.2	Tier-to-tier configuration . . . . .	104
5.2.1	System description . . . . .	104
5.2.2	Queueing network model . . . . .	111
5.2.3	Modeling of the service time distributions . . . . .	116
5.2.4	Computation of the network performance measures . . . . .	121
<b>6</b>	<b>Validation and numerical evaluation . . . . .</b>	<b>129</b>
6.1	Validation . . . . .	129
6.1.1	Description of the discrete-event simulation . . . . .	129
6.1.2	Evaluation of the approximation quality . . . . .	133
6.2	Numerical evaluation . . . . .	151
6.2.1	Design of SBS/RSSs under multiple constraints . . . . .	153
6.2.2	Impact of multi-level shuttles . . . . .	158
6.2.3	Impact of re-entrant bins and variability of random variables . . . . .	161
<b>7</b>	<b>Conclusion . . . . .</b>	<b>171</b>
7.1	Summary . . . . .	171
7.2	Outlook . . . . .	174
<b>Glossary of Notation . . . . .</b>	<b>177</b>	
<b>References . . . . .</b>	<b>197</b>	
<b>List of Figures . . . . .</b>	<b>205</b>	

<b>List of Tables</b>	<b>209</b>
<b>A Discrete distributions and conveyor data of the validation and numerical evaluation</b>	<b>211</b>
A.1 Discrete distributions	211
A.2 Distances of the conveyor system used in the simulation	214
<b>B Application example of a tier-captive system: additional data</b>	<b>215</b>
<b>C Application example of a tier-to-tier system: additional data</b>	<b>219</b>



# 1 Introduction

Uncertainty in demand and supply, production in lots, smoothing of peak demand, and many other reasons require to hold inventory in a warehouse. During the design phase of a warehouse, decisions have to be made about the type of storage system and the equipment which is used for the storage and retrieval process of the goods.

Simple storage types are ground block storage and different kinds of rack storage systems that are manually operated. In these systems, the operator moves to the storage location to store or retrieve the goods. An advantage of simple storage types is that the number of operators, and therefore the achievable throughput, is easy to adapt to the current needs. Also, the investments that have to be made are rather low. On the other hand, the operator can only reach a certain height of a storage shelf or a pallet rack that is served by forklift trucks. This causes inefficiencies in the usage of floor space and height. Moreover, due to the movement of the operators to the goods, the traveling times of the operators represent a considerable share of the total picking time. This causes inefficiencies in throughput as well.

As a result, automated storage and retrieval systems (AS/RSs) were developed to reduce these inefficiencies. In AS/RSs, the goods are moved from their storage locations to the operators by automated material handling equipment. In the last decades, the most common equipment to automatically store and retrieve goods were so-called storage and retrieval machines (SRMs), which are running through aisles between the racks. With SRMs, it is possible to automatically serve high rack storage systems that are up to 55 meters high (ten Hompel et al. (2007)). This yields savings in labor costs and floor space. Disadvantages are less flexibility and higher investments in equipment and control systems (Roodbergen and Vis (2009)). Moreover, especially in mini-load systems, the mass of an SRM is very large in comparison to the mass of the

transported bin. Hence, the energy consumption to store and retrieve a single bin is quite high.

A special type of an AS/RS is a system in which vehicles travel in horizontal direction over rails through aisles and cross-aisles to store and retrieve the goods, while lifts are used for vertical movement. In the literature, such systems are often denoted as autonomous vehicle storage and retrieval systems (AVS/RSs) (Roodbergen and Vis (2009)). In comparison to traditional AS/RSs with SRMs, the throughput of AVS/RSs can be adjusted by adding vehicles and lifts to the system. AVS/RSs for mini-loads with aisle-captive vehicles that cannot change the aisles are often called shuttle-based storage and retrieval systems (SBS/RSs) in the literature. A common usage of SBS/RSs is the retrieval of goods for order picking. In these systems, the SBS/RS is usually connected via a conveyor system to the picking stations. After the picking process, the non-empty bins re-enter the SBS/RS. Hence, the connection between the SBS/RS and the picking stations is an important part of the whole system.

In many SBS/RSs, tier-captive vehicles are installed. In these systems, each vehicle is assigned to exactly one tier of an aisle. Due to the independence of horizontal and vertical movement, tier-captive SBS/RSs reach a high performance (Marchet et al. (2013)) and are installed more and more frequently. A new development in this field is the installation of SBS/RSs with vehicles that are able to serve more than one level of a tier. These multi-level (ML) shuttles<sup>1</sup> increase the utilization of the vehicles due to the fact that they have to serve more storage locations. Hence, the usage of multi-level shuttles can be beneficial in systems where single-level (SL) shuttles would have a low utilization.

Based on the characteristics of the goods as well as the required storage capacity, throughput, and the service level in terms of the retrieval transaction time, the decision maker decides on the system design. Since the system design has a direct impact on the investment costs, operational costs, and penalty costs (if the service level is not met), the decision maker also contributes to the prof-

---

<sup>1</sup> In the literature, the vehicles of SBS/RSs are usually called shuttles. Since they represent the main resources to store and retrieve the goods, the whole system is called SBS/RS. In the later parts of this work, we will use the terms shuttle and vehicle synonymously. Thus, we denote vehicles that are able to serve multiple levels of a tier as multi-level shuttles (sometimes also denoted as *small SRMs* in the literature).

itability of the company. As a result, there is a need for performance evaluation tools that enable the decision maker to evaluate a large number of different system designs in a short period of time. Since the system behavior, and therefore the system performance, is influenced by stochastic processes such as the arrival process of the retrieval transactions, the performance evaluation tools have to consider these influences.

## 1.1 Problem description

The performance evaluation of processes that are subject to stochastic influences can be done either by simulation or by analytical methods. Simulation offers the possibility to analyze the processes in any desired level of detail. This is why simulation is used in the advanced design stages of warehousing processes to predict the system performance. If the design fulfills the required performance, the material handling provider will continue to the next phase, installing the equipment. If the performance is insufficient, the system designer will make final adjustments to overcome the insufficiency. However, the development of simulation models that represent the real system in the needed detail and the execution of the simulation runs is very time consuming. To obtain a steady-state solution, long simulation times are required. Moreover, due to the inexactness of the simulation solution, multiple simulation runs and statistical analysis for the determination of the confidence intervals have to be performed. This makes the development and the execution of the simulation model very costly. Hence, simulation is not suitable for the early planning stages of warehousing systems.

Analytical methods, on the other side, are capable of comparing a large number of different system designs in a short period of time. If the stochastic influences lead to waiting times for resources, queueing models can be used to describe the system behavior. Thus, by representing the system as a queueing system or queueing network, the system designer is able to compute the system performance measures. Moreover, in some cases it is possible to use the outcome of the queueing system as input for optimization. Although the level of detail is lower in comparison to simulation, the results are accurate enough for the early planning stages. Another advantage of queueing models is the gain of general system insights during the modeling and the analysis phase.

In queueing theory, there is the possibility to model the system in the continuous or discrete time domain. Advantages of continuous-time queueing models are short computation times and a set of closed form solutions that can be used for optimization. A disadvantage is the need of distribution functions that describe the stochastic processes. In practice, many stochastic processes cannot be described by simple continuous distributions. For example, processes in which the processing times can only assume certain discrete values or processes that are distributed according to a multi-modal distribution. In these cases, the processing times have to be approximated by complex expressions. Moreover, the complete probability distributions of the performance measures can only be obtained for processes that are described by simple probability functions or a combination of these functions.

In comparison to queueing models in the continuous time domain, discrete-time queueing models can directly use the data obtained by an as-is analysis or the distributions obtained by the evaluation of the physical processes as input for the modeling of the processing times. Furthermore, it allows for the computation of the complete probability distributions of the performance measures. This property is especially of interest if we want to assure a given service level, i.e. that a given proportion of the retrieval transactions leaves the system within a given time span. It is also of interest, if we want to determine the distribution of the number of customers waiting in front of a resource to dimension the needed buffer capacities. Due to the numerical methods that are used to solve discrete-time queueing models, the computation times are generally higher compared to the ones of continuous-time methods. However, the level of detail that can be modeled is higher (Schleyer (2007)). As a result, discrete-time queueing theory represents a reasonable method for the early planning stages of systems under stochastic influences. In particular, if the knowledge about the quantiles of the distributions of the performance measures is of interest.

There is a rich literature on analytical methods for the performance evaluation of manually operated storage systems, AS/RSs, and SBS/RSs. However, regarding SBS/RSs, the existing performance evaluation tools either compute just the average values of the performance measures or consider only specific parts of the system such as the aisle of an SBS/RS, neglecting the connection to the picking stations. An integrated approach that considers both the computation of the quantiles of the performance measures and the modeling of

picking stations and re-entrant bins is still missing. Moreover, investigations on multi-level shuttles in such a system configuration are missing as well.

Hence, the goal of this work is to develop a modeling approach in the discrete time domain that enables the system designer to compute the distributions of the performance measures of SBS/RSSs with multi-level shuttles, picking stations and re-entrant bins. Given this approach, the system designer will be able to design SBS/RSSs given the following system requirements: storage capacity, operational throughput, maximum warehouse size, and service level in terms of a given quantile of the retrieval transaction time distribution. Moreover, we show the impact of multi-level shuttles, re-entrant bins, and the variability of the random variables on the design decision. The random variables of interest are the retrieval inter-arrival time, the storage replenishment inter-arrival time, and the picking time.

## 1.2 Organization of the thesis

The overall structure of the thesis is depicted in figure 1.1. The chapters 1 to 3 are dedicated to the problem description and the exposition of SBS/RSSs. After the introduction, chapter 2 focuses on the characterization of SBS/RSSs. Based on a high level system description including the flow of the storage and retrieval transactions, we describe the different physical designs and control policies that characterize an SBS/RSS. In addition, we present the typical design requirements and the performance measures which are used to identify suitable system designs.

In chapter 3, we give a review of analytical and simulation models that are available for performance evaluation of SBS/RSSs. We conclude this chapter with a summary of the literature review, and show the need for an integrated approach for the performance evaluation of SBS/RSSs based on queueing models in the discrete time domain.

In the following chapters 4 to 5, we present the discrete-time models of SBS/RSSs. In chapter 4, we describe the general modeling and solution approach for the performance evaluation of SBS/RSSs. It is based on the decomposition of the discrete-time queueing network into  $G|G|1$  queueing systems. Therefore, we first introduce the basics of discrete-time probability theory as

well as the models and methods in the discrete time domain that are used in the modeling approach.

In chapter 5, both tier-captive and tier-to-tier SBS/RS configurations are modeled using the general modeling approach. Based on detailed system descriptions, we model the systems as open queueing networks, and model the service time distributions by discrete probability distributions that are obtained by the analysis of the system dynamics. Afterward, we decompose the network into  $G|G|1$  queueing systems and use the discrete-time methods to compute the network performance measures.

Following the presentation of the modeling approach, we evaluate the approximation quality of the decomposition approach by comparison to discrete-event simulation (DES) in chapter 6. Moreover, we perform a numerical evaluation to show how the system designer can use the approach for decision making during the design phase of SBS/RSs. In addition, we show the impact of multi-level shuttles, re-entrant bins and the variability of the random variables on the design decision.

Finally, chapter 7 summarizes the contribution and main results of this work, and gives an outlook on future research.

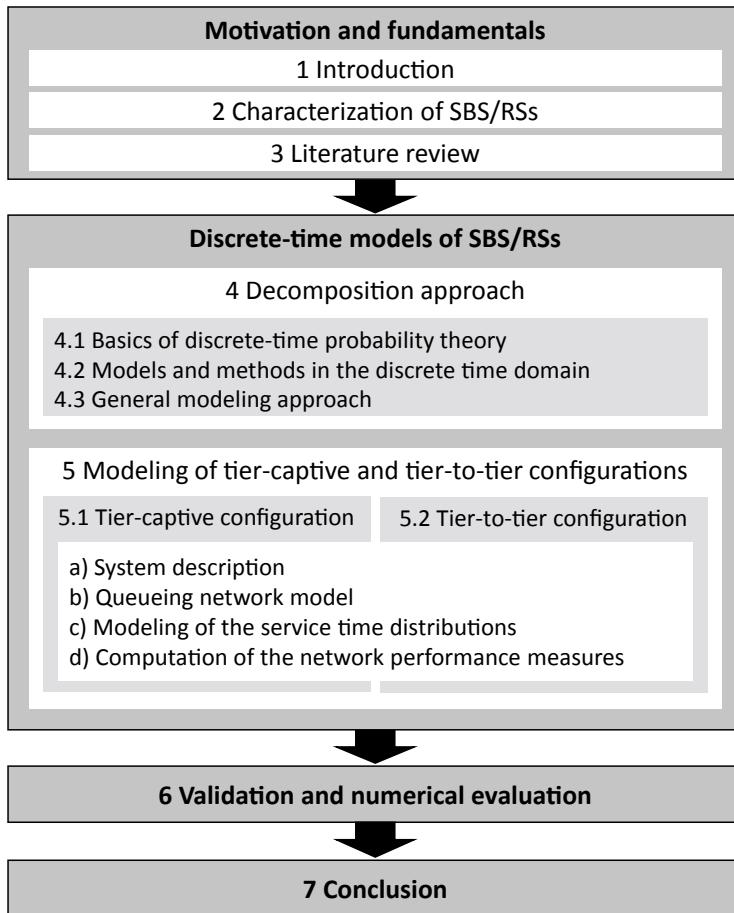


Figure 1.1: Outline of the thesis



## 2 Characterization of SBS/RSs

Shuttle-based storage and retrieval systems are installed more and more frequently in warehouses all over Europe. They represent a special type of automated storage and retrieval system for mini-loads, in which vehicles are used to store and retrieve goods that are stored in containers, bins, or trays within the rack system. The vehicles travel horizontally along the aisles and cross-aisles of the rectangular-shaped rack system and store/retrieve the goods using a load handling device (LHD). The vertical movement of the goods between the tiers and the input/output (I/O) points of the SBS/RS usually is done by using lifts. From the output point of the SBS/RS, the goods are transported to the subsequent processes. Since SBS/RSs are often used in part-to-picker systems, conveyor systems usually transport the goods to the nearby picking stations. At the picking stations, operators pick the needed amount of parts out of the bin. Afterward, the conveyor system transports the non-empty bins back to the input point of the SBS/RS and the empty bins out of the SBS/RS system boundaries for replenishment. Figure 2.1 depicts the typical position of an SBS/RS within the material flow of a warehouse.

Automated storage and retrieval systems, in which vehicles are used to store and retrieve goods, can be classified based on the degree of freedom (DoF) of the vehicles. If the vehicles are able to change tiers by using the lift, they are referred to as tier-to-tier vehicles. If the vehicles are not able to change the tiers, which means that the lift only transports the goods, they are referred to as tier-captive vehicles. As a result, we use the terms tier-to-tier configuration and tier-captive configuration to denote these kinds of systems. The same applies to the possibility of changing the aisle. Vehicles that are able to change the aisle by using cross-aisles are referred to as aisle-to-aisle vehicles, and vehicles that are not able to change the aisle are referred to as aisle-captive vehicles. An additional differentiation can be made by the number of levels in the storage rack that a vehicle serves. If the vehicles are able to serve more than one level,

we refer to them as multi-level shuttles. If they are only serving one level of a storage rack, we refer to them as single-level shuttles<sup>1</sup>.

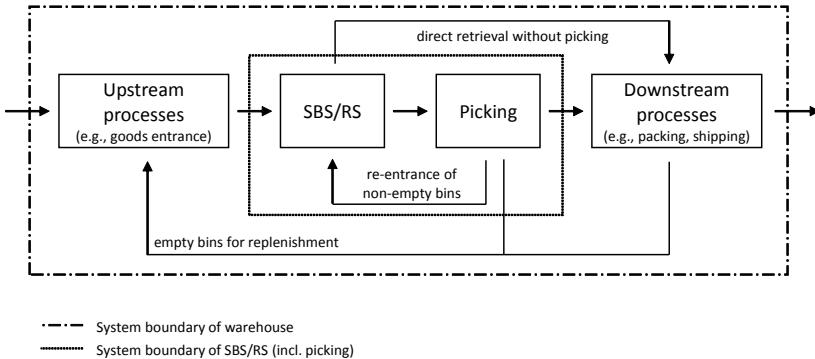


Figure 2.1: Position of the SBS/RS within the material flow of a warehouse

In the literature, systems with a high DoF of the vehicles are usually denoted as AVS/RSSs, whereas systems with a low DoF of the vehicles are often called SBS/RSSs (see also figure 2.2). Moreover, the literature on AVS/RSSs is mainly analyzing storage systems for unit-loads, whereas the literature on SBS/RSSs is often assuming mini-loads. Indeed, both notations can be used to describe the same system. That is why there are publications on systems with a high DoF of the vehicles, in which the system is called SBS/RS (and the other way around). Since we want to focus on systems for mini-loads with a rather low DoF of the vehicles, we denote the systems under investigation in this thesis as SBS/RSSs.

As stated in the introduction, there are several advantages of SBS/RSSs over traditional AS/RSSs for mini-loads. Given a tier-captive configuration, the mutual independence between the vehicle and lift movement yields a high performance in terms of throughput (Marchet et al. (2012)). In terms of scalability, using a tier-to-tier configuration results in a high flexibility since additional vehicles can be added without changing the rack configuration according to

<sup>1</sup> As stated in the introduction, a differentiation between these systems can also be made by denoting ML shuttles as *small SRMs* and SL shuttles as *shuttles*.

match the needed throughput (Malmborg (2003)). Moreover, Schmidt (2010) states that the high redundancy due to the large number of vehicles increases the availability of the system. He also states that the movement of relative low masses reduces the need for energy. In addition, he points out that the usage of SBS/RSs makes it possible to sequence the bins in the right order.

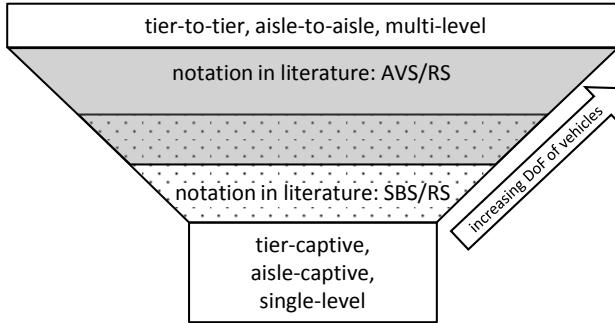


Figure 2.2: Different notations in literature based on the DoF of the vehicles

However, the large number of material handling equipment used in SBS/RSs results in complex system designs. Regarding the rack system, the system designer needs to include a rail system on each tier such that the vehicles are able to reach the storage locations. Depending on the system configuration, the tiers may also have to contain an energy supply system and a communication system. This leads to a more sophisticated rack design. Regarding the system control, the large number of parallel resources increases the complexity of the control strategies. Since the physical design and the control strategies both influence the system performance, the estimation of the performance measures and, therefore, the decision on the system design becomes more complex.

In this work, we develop a modeling approach that can be used to analyze SBS/RSs under different system configurations. This approach makes it possible to design SBS/RSs such that the design fulfills the demand requirements while avoiding bottlenecks and overcapacity. In order to create a better un-

derstanding of the modeling approach, in this chapter, we briefly present the possible system designs of SBS/RSs.

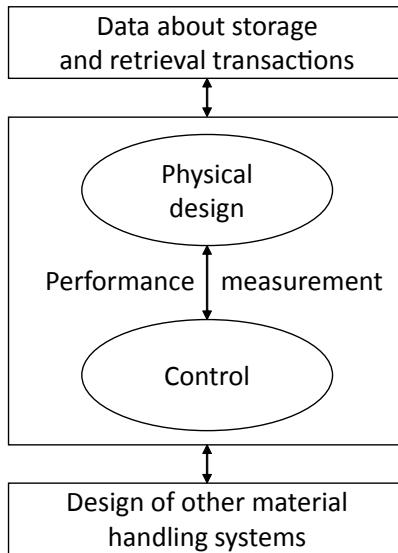


Figure 2.3: SBS/RS design (based on the AS/RS design of Roodbergen and Vis (2009))

According to Roodbergen and Vis (2009), the design of an AS/RS is defined by the physical design and the control strategies applied (see figure 2.3). Moreover, it is influenced by the design of other material handling systems and the arrival of storage and retrieval transactions. Therefore, we will first give a description of the flow of storage and retrieval transactions (section 2.1). Afterward, we discuss possible physical design configurations (section 2.2) and control strategies (section 2.3). In addition, we present common design requirements and performance measures of SBS/RSs that have to be taken into consideration when designing an SBS/RS (section 2.4).

## 2.1 System load – flow of storage and retrieval transactions

The system load on the resources of the SBS/RS consists of storage and retrieval transactions, which represent individual transport orders of the warehouse management system. The transactions are arriving and/or leaving the resources either virtually in form of information or physically in form of a bin. For example, the transport of a bin, in which an item of a picking order is stored, from its storage location in the SBS/RS to the picking station is generating a retrieval transaction. This transaction arrives virtually in form of information at the SBS/RS and leaves the SBS/RS physically in form of a bin. Also the transport of a bin from the SBS/RS to other subsequent processes than picking and the transport of empty bins from the picking station to the stations for replenishment are generating retrieval transactions. In contrast, the transport of a non-empty bin from a picking station back to the SBS/RS is generating a storage transaction. The transport of bins for replenishment from preceding processes to the SBS/RS is generating storage transactions as well.

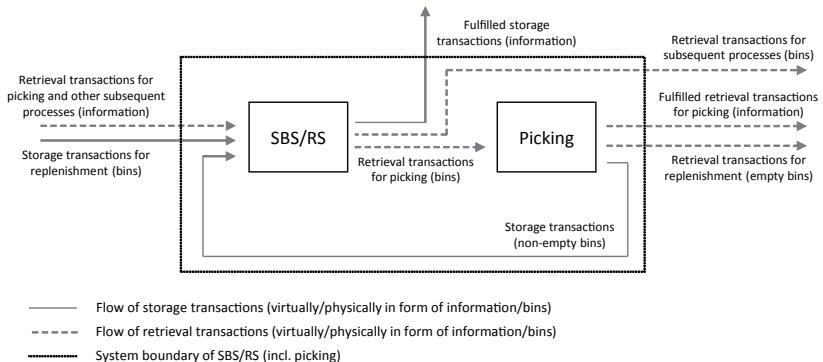


Figure 2.4: System load – flow of storage and retrieval transactions

The flow of the storage and retrieval transactions in and out of an SBS/RS that is used in a part-to-picker system is depicted in figure 2.4. Since there is a frequent interaction between the SBS/RS and the picking process, we also show

the flow of transactions in between the SBS/RS and the picking process. We see that the generated retrieval transactions for picking and other subsequent processes virtually arrive at the SBS/RS and induce a load on the resources of the SBS/RS that are used to transport the bins out of the SBS/RS (mainly vehicles and lifts). After the transport out of the SBS/RS, the retrieval transactions (bins) for subsequent processes other than picking use the conveyor system to reach these processes, whereas the retrieval transactions (bins) for picking use the conveyor system to reach the picking stations (pickers). After the picking process, the retrieval transactions for picking are fulfilled and virtually leave the network. If a bin was emptied by a picker, the transport of the empty bin to the station for replenishment is generating a new retrieval transaction that leaves the network via the conveyor system. If a bin was not emptied by a picker, the transport of the non-empty bin from the picking station via the conveyor system to its storage location within the SBS/RS is generating a storage transaction which is inducing a load on the resources of the SBS/RS that are used to transport the bin into the SBS/RS (mainly vehicles and lifts). The same applies for the transport of a bin, whose items have been replenished, to its storage location within the SBS/RS. After reaching the storage location, the storage transactions are fulfilled and virtually leave the network. In systems without a picking process, the flow of storage and retrieval transactions is only consisting of the generated storage and retrieval transactions that are inducing a load on the SBS/RS, and the flow of storage and retrieval transactions (virtually) leaving the SBS/RS.

In general, this results in an open flow of storage and retrieval transactions through a network of resources inside and outside of the SBS/RS that are connected to the other parts of the warehouse via conveying systems. The transactions enter the network virtually in the case of retrieval transactions, and physically via a conveying system in the case of storage transactions for replenishment. They leave the network virtually in the case of the fulfilled storage transactions and the fulfilled retrieval transactions for picking, and physically in the case of the retrieval transactions for subsequent processes and the generated retrieval transactions of empty bins for replenishment.

The resources that are used by the transactions inside and outside of the SBS/RS could be both, active conveying systems such as roller conveyors, and passive conveying systems (e.g., rails) in combination with discontinuous con-

veying devices (e.g., shuttles). Due to the limited number and capacity of the conveying systems and devices, there are parts of the network with a population constraint, i.e. only a specific number of transactions is allowed to use the given part of the network at the same time. Based on the source and destination of the transport order, a transaction uses multiple resources during its time in the system. Hence, every source/destination combination represents an individual customer class and has its own deterministic route through the network of resources.

The generation of retrieval transactions for picking and other subsequent processes creates inter-arrival times that are distributed according to a non-negative continuous random variable. The same applies for the generation of storage transactions for replenishment, except that the minimum distance between two bins on a conveying system leads to a lower bound of the inter-arrival times that is larger than zero. Within the network of resources, the inter-arrival times at the subsequent resources depend on the inter-departure times of the preceding resources and the routing of the transactions.

The processing time of a transaction on a specific resource of the network is state-dependent. It depends on the current location of the resource, the locations of the source and destination of the transaction, and the transfer (loading/unloading) times, velocities and acceleration/deceleration rates of the material handling device. Due to the finite number of locations and the deterministic transfer times, velocities and acceleration/deceleration rates, the processing times of a transaction on a resource may only assume a finite number of discrete values. At the picking stations, the processing times are dependent on the number and characteristics (size, weight, etc.) of the picked items as well as the performance of the picker. They may be distributed according to a non-negative continuous random variable.

## 2.2 Physical design

In this section, we give a brief description of different system designs of SBS/RSs. Since many SBS/RSs are directly connected to picking systems, we also include the picking process and possible conveying systems in between the SBS/RS and the picking process. An example of a tier-captive SBS/RS

configuration, which is connected via a conveyor system to a picking station, is depicted in figure 2.5.

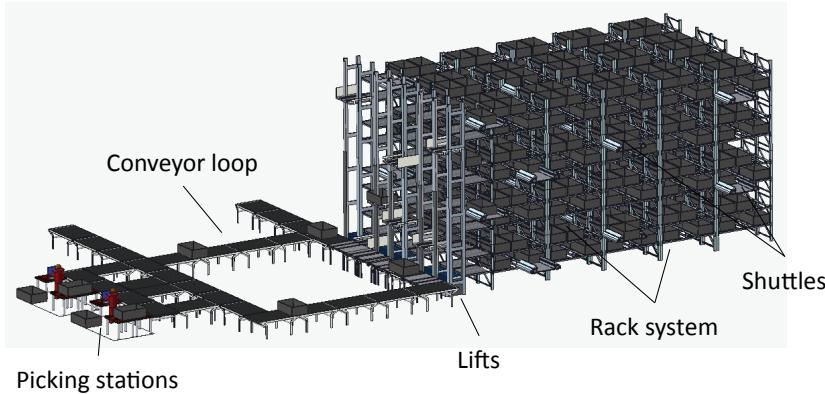


Figure 2.5: Example of an SBS/RS with a conveyor loop and two picking stations

In general, the physical design of an SBS/RS consists of primary and secondary resources that are used to store and retrieve the goods to and from the SBS/RS. The primary resources are the rack system as well as the active and passive conveying systems (e.g., roller conveyors, rails) inside and outside of the SBS/RS. The secondary resources are the discontinuous conveying devices (e.g., shuttle, lift) that are using the passive conveying systems, i.e. primary resources, to transport the goods. An overview of the physical design and its primary and secondary resources is given in figure 2.6. The dotted lines indicate possible connections between the primary and secondary resources.

Furthermore, the design of an SBS/RS also includes systems for communication and energy supply. For a detailed description of the physical design, in particular the systems for communication and energy supply, the interested reader is referred to VDI-Richtlinie 2692 (2015).

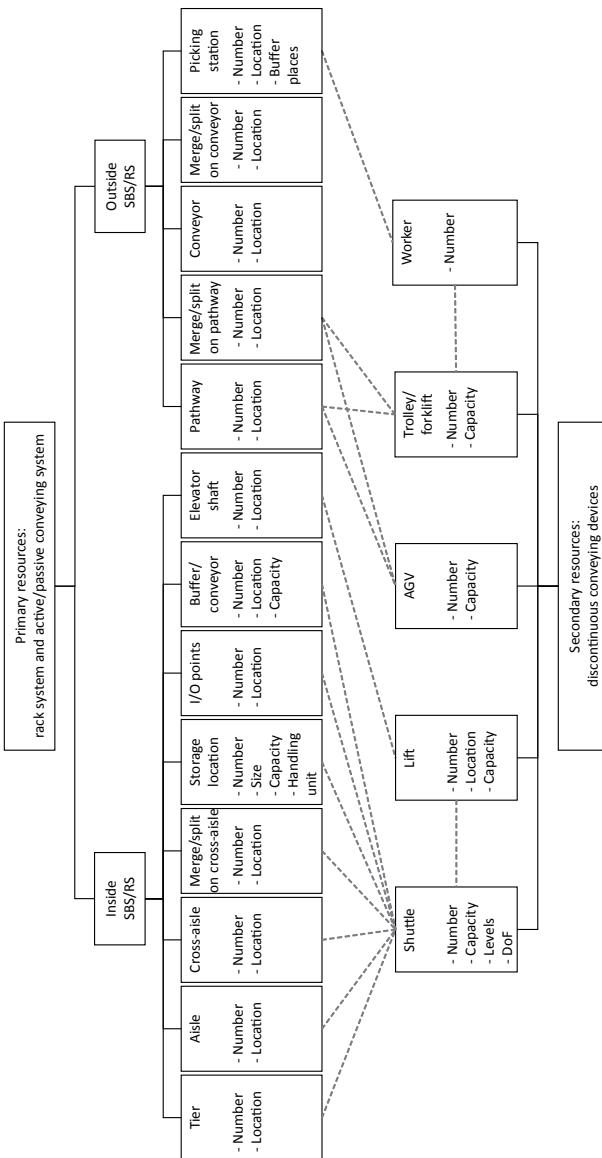


Figure 2.6: Physical design

## Primary resources

The rack system of an SBS/RS is mainly defined by the number and location of the tiers, aisles, and cross-aisles, as well as the number, size, and capacity of the storage locations. The rack itself is more or less identical to the rack used in AS/RSs for mini-loads, except that there are rails along the aisles and cross-aisles of each tier which are used by the vehicles as passive conveying systems. The handling units that are stored in the storage locations are usually trays, bins, or containers, since they are easy to handle by the load handling device of the vehicle. Depending on the need for a direct and fast access to the goods, the weight of the handling units, and the possible savings in floor space, the capacity of the storage locations can vary between single-deep, double-deep, and multi-deep storage. The size of the handling unit and the capacity of a storage location determine the storage locations' size. Depending on the size of the goods and the possibility of the vehicles and lifts to handle different sized bins or trays, equally sized or modular storage locations may be installed.

Given the number and size of the needed storage locations as well as the maximum floor space and height of the system, the system designer is able to determine the possible combination of number and location of tiers and aisles as well as the length of the aisles. The existence of cross-aisles is dependent on whether the vehicles are able to change the aisles. Since the cross-aisles have an impact on the traveling time of the vehicles, the determination of their number and location is an important design decision. If cross-aisles exist, they determine the number and location of the intersections with the aisles, where the material flow can merge and split. The intersections may be passive conveying elements or active conveying elements such as turntables.

The elevator shafts that are used by the lifts as passive conveying systems are also part of the SBS/RS. In aisle-captive systems, at least one elevator shaft per aisle is needed to transport the goods of this aisle into or out of the SBS/RS. Depending on the requirements for throughput, also more than one elevator shaft per aisle may be installed. If more than one shaft is installed, often one shaft is assigned to storage transactions and the other one to retrieval transactions. The location of the elevator shafts usually depends on the connection of the SBS/RS to the adjacent conveying system. In many cases, they are located at the beginning of an aisle since the conveying system is just connected to one side of the SBS/RS. In aisle-to-aisle systems, the number and location

of the elevator shafts is more flexible since the vehicles of a tier can reach every position within the tier. Hence, in such systems it is important to determine the appropriate number and location of elevator shafts to reach the needed throughput. Depending on the transport of only the goods or the vehicles and the goods, the sizes of the elevator shafts vary.

In tier-captive systems, buffer places (usually active conveying systems) are installed at every tier to create a mutual independence between the vehicles and the lifts. Given retrieval transactions, the vehicles unload the goods onto the outgoing buffers, from where they are transported to the point where the lift picks up the outgoing goods. Given storage transactions, the lift unloads the goods onto the incoming buffers, from where they are transported to the point where the vehicles pick up the incoming goods. Thus, they are located between the tier and the elevator shaft. Since the number of buffer places has a direct impact on the probability of blocking, the dimensioning of those places is an important design decision. The same is valid for the buffer places (active conveying systems) between the I/O points and the connected conveying system.

The input and output points of the SBS/RS represent the interface of the SBS/RS to the connected conveying system. At these points, the goods are entering or leaving the SBS/RS. The input point is the place where the incoming goods are transferred from the buffer to the lift. At the output point, the outgoing goods are transferred from the lift to the buffer between the output point and the conveying system. Hence, the number and location of the I/O points depend on the number and location of the elevator shafts. The vertical locations of the input and output points can be different. As a result, the determination of their locations is another design decision since they influence the travel times of the lifts between the I/O points and the tiers.

Outside the SBS/RS, active and passive conveying systems may be used to transport the goods to the input points and from the output points to the subsequent processes. In most cases, there is an active conveying system connected to the SBS/RS. It consists of roller and/or belt conveyors that transport goods to and from the SBS/RS. The flow of goods determines the layout of the conveyor system. The basic elements of the conveyor system are conveyor lines and intersections, where the material flow can merge and/or split. If there is no active conveying system connected to the SBS/RS, pathways (with/without rails) can be used by automated guided vehicles (AGV) or workers with fork-

lifts/trolleys to transport the goods. Different directions of the material flow results in intersections, i.e. merges and splits on the pathways/rails.

The picking process is often the subsequent process after the retrieval of the goods. At the picking station, a picker is taking the needed items out of the bins and places them into a new container or carton box. Based on the time for the picking process and the number of picks that have to be fulfilled, the number of picking stations is determined. After picking, the non-empty bins are routed back on the conveying system to the SBS/RS for storage, whereas the empty bins are leaving the system for replenishment. The packed new containers or carton boxes are usually leaving the picking station manually via forklifts or trolleys, or via another conveyor system. Since the bins should not move while the items are being picked, there are split and merge elements that direct the bins from the main conveyor line to the designated picking stations and back. In front of the picking stations, the variability of the inter-arrival time of the bins that are routed to the picking stations and the variability of the service times at the picking stations create a waiting process. As a result, there are buffer places (active conveying systems) in front of the picking places. To create the mutual independence between the picking stations and the flow of goods on the main conveyor line to the different picking stations, the dimensioning of the number of buffer places in front of the picking stations has to be done carefully.

### **Secondary resources**

Inside the SBS/RS, vehicles and lifts are used as secondary resources. They are discontinuous conveying devices that are using the primary resources to transport the goods simultaneously and independently from each other in horizontal and vertical direction. Vehicles are used to transport the bins in horizontal direction, and to store and retrieve the bins to and from the rack. As stated before, these vehicles are usually called shuttles.

Different types of shuttles can be installed in SBS/RSs. They are classified based on their degree of freedom. In most applications, they are only capable of traveling in a single direction along the rails of an aisle, which makes them aisle-captive shuttles. If they are able to move in two directions by traveling on cross-aisles, they are called aisle-to-aisle vehicles. The movement in two

directions can be achieved by rotatable wheels. But also other techniques such as a turntable at the intersection of an aisle and a cross-aisle can be used to change the direction of a shuttle.

Another degree of freedom is the possibility of changing the tiers by using the lift. This results in the differentiation between tier-captive and tier-to-tier vehicles. In tier-captive configurations, at least one vehicle is needed in every tier. If the vehicles are also aisle-captive, one vehicle is needed in every aisle of a tier. This results in short traveling times and therefore a high throughput. In tier-to-tier configurations, the number of vehicles can be adjusted according to the current needs. This makes the tier-to-tier configuration more flexible in comparison to the tier-captive configuration.

To store and retrieve the goods to and from the rack, load handling devices are installed on the shuttles. Different load handling techniques such as gripping devices, belt conveyors, and clamping devices can be used to store and retrieve the goods. Depending on the used technique, differently or just equally sized bins or trays can be loaded to the shuttle. The capacity of one LHD can vary between a single bin and multiple bins. The number of LHDs per vehicle and their capacity yields the capacity of the vehicle.

As stated in the introduction, developments have resulted in shuttles, where the LHD can reach more than one level of the rack system. They have an additional lifting system for the LHD that makes it possible to reach more than one level of the rack system without changing the tiers. This increases the DoF of the shuttle/LHD. The shuttles that are capable of this we denote as multi-level shuttles, whereas the shuttles that can only reach one level of the rack system we denote as single-level shuttles. The lifting of the LHD can be done simultaneously to the horizontal movement of the shuttle, which results in the same movement pattern as the one of traditional AS/RSs. Theoretically, if there is just one vehicle installed that can reach all levels of the rack system, its movement will be identical to the movement of an AS/RS<sup>2</sup>.

To transport the bins in vertical direction, usually lifts are used as secondary resources. They use the elevator shafts as passive conveying element to transfer the bins with or without the shuttles between the I/O points and the tiers of the

---

<sup>2</sup> That is why ML shuttles are sometimes also denoted as *small SRMs* in the literature.

SBS/RS. Hence, they are located inside the elevator shafts. The number of lifts in the system depends on the possibility to install only one or multiple lifts in a shaft. Each lift usually has one platform, on which a conveying technology to transport the bins to or from the platform is installed in the case of tier-captive vehicles. In tier-to-tier configurations, a rail is installed on the platform such that the vehicles can directly travel onto the platform. Thus, the capacity of the lift depends on the number of platforms and the capacity (number of bins/vehicles) per platform. Other secondary resources that can be used for the vertical transport of the goods are circulating vertical conveyor systems that consist of multiple circulating platforms.

If there is no active conveying system installed outside the SBS/RS, secondary resources are needed to transport the goods on passive conveying systems to and from the SBS/RS. In the few cases, where no conveying system is installed, different discontinuous conveying devices can be used. For example, automated guided vehicles or automated forklift trucks are able to take over the transport of the bins to and from the SBS/RS. But also workers with manually operated trolleys or forklifts can be used to transport the goods. At the picking stations, workers are used as secondary resources to pick the needed items out of the bins.

## 2.3 Control policies

In addition to the physical design, the applied control policies have an important impact on the system behavior and therefore on the performance measures of the AS/RS. Frequently used and analyzed control policies of AS/RSs are described in detail by Roodbergen and Vis (2009). Since SBS/RSs are a special type of AS/RSs, similar control policies can be applied. Based on the survey of Roodbergen and Vis (2009), we describe in this section the control policies that can be applied to SBS/RSs.

The control policies for SBS/RSs can be classified by five main subjects: storage assignment, sequencing, resource assignment, routing, and dwell point strategy (see also figure 2.7). Whereas storage assignment, sequencing, and dwell point strategies are well studied design decisions in AS/RSs (see Roodbergen and Vis (2009)), resource assignment and routing strategies are less

studied. However, due to the possibility of multiple parallel resources and different paths through a network of aisles and cross-aisles, we add resource assignment and routing strategies to the control policies of SBS/RSs. We do not consider batching, since both the determination of the batch size and the combination of orders to batches are questions that only have to be answered in person-on-board item-picking AS/RSs and manually operated picking systems. The question regarding the combination of orders to multi command cycles will be answered in the sequencing strategy.

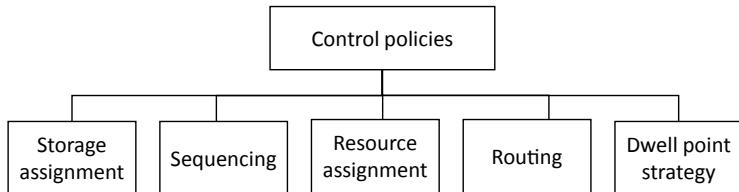


Figure 2.7: Control policies (extension of the AS/RS policies of Roodbergen and Vis (2009))

### **Storage assignment**

The storage assignment strategy determines the storage location of a storage transaction. Frequently used strategies are dedicated, random, closest open location, turnover-based, and class-based storage assignment. In a dedicated storage assignment strategy, every product has its fixed number of dedicated storage locations. This strategy might be applied in retail stores, in systems that cannot keep track of the storage locations of the products, or in systems where heavy products have to be stored in dedicated locations. A disadvantage is the reservation of a fixed storage capacity over a long period of time, even if the products are out of stock.

To generate pooling effects, random storage assignment can be applied. In this strategy, an incoming storage transaction is assigned to all empty storage locations with the same probability. If the inventory levels of all products are not perfectly positive correlated, this results in a lower number of total storage locations.

In order to reduce travel times for storage transactions, the closet open location strategy can be applied. In this strategy, a storage transaction is stored to the first empty location that is reached by the material handling equipment. This results in full storage locations around the I/O points. Since also products with a low turnover value are stored around the I/O points, this strategy may cause long travel times for products with a high turnover value that could not be stored close to the I/O points.

To overcome this disadvantage, the turnover-based storage assignment strategy can be used. In this strategy, the products with the highest turnover value are assigned to the locations that are reached first by the material handling equipment. Usually, these are the locations closest to the I/O points. However, since the demand frequencies change over time and new products are added to the AS/RS, a large amount of repositioning activities may be needed.

To combine the benefits of random storage, which does not require repositioning, and dedicated storage based on the turnover value, class-based storage can be applied. In this strategy, the storage locations are divided into different areas. Every product is assigned to one of the areas based on its turnover value. On the basis of an ABC classification, in many cases three areas for fastest-moving (class A), normal-moving (class B), and slowest-moving (class C) items are installed. Within the areas, random storage assignment is applied. In general, the questions to be answered during the design phase are the decision on the number of areas and the number of products that are assigned to each area, as well as the positioning of the areas.

## **Sequencing**

The sequencing strategy of an SBS/RS is defined by the method to determine the sequence of the storage and retrieval transactions that a secondary resource (shuttle/lift) has to fulfill. Before we decide on the sequence, we first have to decide if we dynamically update the sequence every time a new transaction enters the queue or if we fulfill the transactions block-wise. In addition, we need to decide whether we want to operate the system under single or dual command cycles. In a single command cycle of a storage transaction, the resource travels to the loading point, loads the goods, travels to the storage location, and unloads the goods. In analogy, in a single command cycle of a retrieval trans-

action, the resource travels to the retrieval location, loads the goods, travels to the unloading point, and unloads the goods. If two transactions from the same type (storage or retrieval) are sequenced right after each other, this strategy can result in long empty travel times. Hence, dual command cycles can be applied, in which every storage transaction is paired with a retrieval transaction. After the storage transaction is fulfilled, the resource travels to the retrieval location, loads the goods, and returns to the unloading point. Obviously, the average total time to fulfill both a storage and retrieval transaction is reduced. However, in a strict dual command strategy, the resource has to wait for a storage transaction, even if a retrieval transaction is already available. Hence, hybrid modes are often applied, in which a single command is performed if a dual command is not possible.

Subsequently, we need to determine the sequence of single and dual command cycles. Since the general dynamic sequencing problem is hard to solve optimally, a common way is to sequence the waiting transactions based on the first-come-first-serve (FCFS) strategy. Storage transactions that are physically waiting in front of a resource, and that are not able to pass each other, have to be handled FCFS anyways. If the retrieval transactions do not have any due date restrictions, different sequencing strategies are possible. In analogy to the storage transactions, retrieval transactions can be sequenced based on a FCFS rule. But also other methods such as the shortest completion time strategy may be applied. In this strategy, the retrieval transaction with the shortest traveling time will be fulfilled first. Moreover, in dual command cycles, we can pair/sequence the storage and retrieval requests such that the distance between the storage and retrieval locations is minimal. This strategy is denoted as nearest-neighbor strategy.

The sequencing problem obviously becomes more complex if we consider due dates of retrieval transactions, priority rules, aisle-to-aisle configurations, shuttles with multiple load handling devices that can perform multi command cycles, rack systems with double-deep storage locations and multiple I/O points, and the location selection problem in case identical products are stored at multiple locations. According to Roodbergen and Vis (2009), only a few heuristics have been developed for these problems.

In addition, some material handling providers of SBS/RSSs state that it is possible to sequence the transactions without the use of a larger sequencing el-

ement (e.g., storage towers) such that the retrieval transactions of a picking order that includes multiple orderlines arrive in the correct sequence at the picking station. The corresponding sequencing algorithms have not been published though.

## Resource assignment

In traditional AS/RSs for mini-loads, the transactions of an aisle are assigned to one SRM. In contrast, the transactions of an SBS/RS may be assigned to a set of aisle-to-aisle and/or tier-to-tier shuttles that can reach the same storage locations of the SBS/RS. Depending on the transaction type and the location and state (busy/idle) of the vehicles within the storage rack upon request, different vehicle assignment rules may result in more or less beneficial vehicle travel times. The same applies for the assignment of the transaction to one of the lifts, if multiple alternatives are available. Therefore, it is important to assign the transactions to the resources in the right way. As for the storage assignment and sequencing strategies, different policies can be applied. In the literature, analytical models of AVS/RSs and SBS/RSs are usually assuming random or FCFS assignment strategies, since they are easy to model. However, other methods that lead to shorter transaction times are conceivable. For example, a method that considers the empty travel times of the idle vehicles to pick up the goods, or a method that considers the travel times of the vehicles to the lifts. These alternative methods have not been addressed in the literature. Thus, the adoption of methods from resource assignment problems of other systems and the development of SBS/RS-specific methods represents a future field of research.

## Routing

In a configuration, in which the vehicles are not able to change the aisles, the routing problem is of no interest since a shortest path routing strategy will be applied. In a system where the vehicles can change aisles and possibly also tiers, the routing problem becomes a more complex design decision. Obviously, a shortest path can be computed for a single, dual, or multi command cycle. However, the routing problem becomes more complex if multiple vehicles try to use the same aisles and cross-aisles at the same time. Since blocking

and deadlock situations might occur, special routing strategies may be implemented that minimize blocking and prevent deadlocks. Similar to the resource assignment strategies, the literature is short on routing strategies for SBS/RSs. Hence, a future field of research will be the adoption of known routing strategies of similar systems such as AGV systems or manual order picking systems.

### **Dwell point strategy**

Next to the decisions on the aforementioned control policies, the decision on the dwell point strategy represents an important step during the SBS/RS design. If a vehicle or lift is idle after fulfilling a transaction, it is traveling to the so called dwell point. Thus, its location has a significant impact on the travel time of the next transaction. The determination of the dwell point can be static, or dynamic based on the state of the system. Most static strategies are named after the location of the dwell point. Hence, common static strategies in AS/RSs are the point-of-service-completion (POSC) strategy, the return-to-I/O-point (RIO) strategy, and the midpoint strategy. Next to these static dwell point rules, also dynamic dwell point rules have been developed. However, only a few rules have been developed for system configurations other than single unit-load capacity AS/RSs (Roodbergen and Vis (2009)). Therefore, further research has to be conducted on dwell point strategies of other AS/RS configurations, including SBS/RS configurations.

## **2.4 Design requirements and performance measures**

The design requirements and performance measure of AS/RSs that are described by Roodbergen and Vis (2009) can be also applied to SBS/RSs. In general, the system designer has to design the SBS/RS in such a way that it can efficiently handle current and future demand and capacity requirements. In addition, the system design should avoid bottlenecks and overcapacity, and should also consider the impact on the other parts of the warehouse that are connected to the SBS/RS. Given the set of systems that fulfill the design re-

quirements, the system designer may choose the system design that creates the lowest annualized costs, which also include the discounted investment costs.

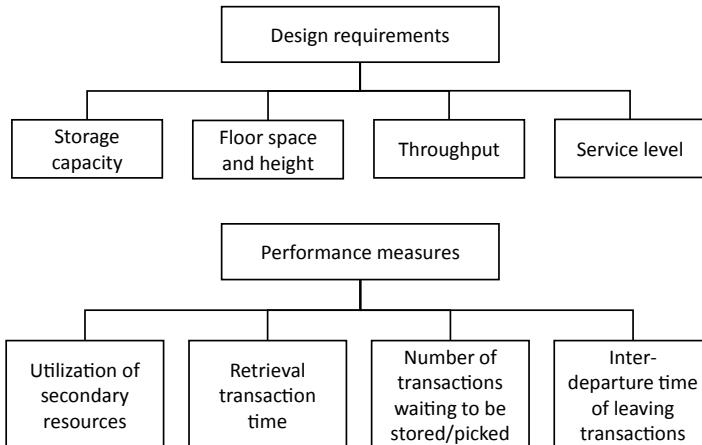


Figure 2.8: Design requirements and performance measures (based on the AS/RS requirements and measures of Roodbergen and Vis (2009))

The major design requirements that arise during the planning phase of an SBS/RS are the following: storage capacity, floor space and height, throughput, and service level (see also figure 2.8). It means that the SBS/RS should be designed in such a way that the storage capacity in terms of the total number of storage locations is identical to or just slightly larger than the needed number of storage locations. Moreover, the floor space and height of the SBS/RS design should not exceed the available floor space and height. Further requirements are the throughput and service level. The throughput is defined as the number of storage and retrieval transactions that have to be fulfilled in a given time period. The service level can be defined by the capability of the SBS/RS to retrieve a predefined proportion (e.g.,  $u\%$ ) of the retrieval transactions within a given time span. The time to retrieve the goods is denoted as the retrieval transaction time. It is the time span from the arrival of the retrieval transaction at the SBS/RS until the departure at the output point of the SBS/RS. In addition, the system design should not negatively influence the other parts of

the SBS/RS. This might be the case if the buffer places in front of the input points of the SBS/RS or the picking stations are not dimensioned sufficiently such that the conveyor loop in front of the SBS/RS is blocked. This can negatively influence the preceding and succeeding processes in terms of blocking and starving. Moreover, the succeeding processes are also influenced by the inter-departure time distribution of the transactions leaving the system.

Based on the number of aisles and tiers as well as the number and capacity of the storage locations on either side of the aisles, we can check whether the system design fulfills the storage capacity requirements. If we know the number of aisles and tiers, the width and length of an aisle, and the height of a tier, we are also able to determine whether the system design fulfills the floor space and height requirements.

Given the inter-arrival time distributions of the incoming storage and retrieval transactions as well as the physical design and the control policies of the SBS/RS, we are able to compute the inter-arrival and service time distributions at the secondary resources, thus the utilization of these resources. Moreover, we can compute the retrieval transaction time distribution, the distribution of the number of transactions waiting to be stored and picked, and the inter-departure time distribution of the transactions leaving the system.

If the utilization of every resource is below 100%, the throughput requirements of the SBS/RS are met. If the  $u\%$  quantile of the retrieval transaction time distribution is smaller than the required time span, the service level of the SBS/RS design is met. If the buffer places in front of the input points of the SBS/RS and the picking stations are dimensioned in such a way that they have a capacity as large as a high quantile (e.g., 99%) of the distribution of the number of transactions waiting to be stored and picked, the other parts of the system are negatively influenced only with a small probability (e.g., 1%). If the inter-departure time distribution of the transactions leaving the system is known, the buffer places of the subsequent processes can be dimensioned accordingly.

As a result, the performance measures of interest are the utilization of every secondary resource, the retrieval transaction time distribution, the distribution of the number of transactions waiting to be stored and picked, and the inter-departure time distribution of the transactions leaving the system (see also figure 2.8).



# 3 Literature review

In this chapter, we give a review on models estimating the performance measures of SBS/RSSs and AVS/RSSs. In the literature, SBS/RSSs with a high DoF of the vehicles are denoted as AVS/RSSs. If we assume that an AVS/RSS consists of a single aisle, these models can also be applied to analyze SBS/RSSs. Thus, we first present publications on analytical and simulation models of AVS/RSSs. Afterward, we focus on existing analytical and simulation models of SBS/RSSs.

## 3.1 Publications on AVS/RSSs

Since this work focuses on analytical performance evaluation models, we will first present the existing analytical models of AVS/RSSs. An overview of these models is given in table 3.1.

### 3.1.1 Analytical models

The first investigations on AVS/RSSs are conducted by Malmborg (2002). He analyzes an AVS/RSS with tier-to-tier vehicle movement, a random storage policy as well as single and dual command cycles. For estimating individual cycle time components, he develops separate material flow matrices for vehicles and lifts.

Using a state equation model, Malmborg (2003) extends his existing model for estimating the proportion of dual command cycles of tier-to-tier AVS/RSSs with opportunistic interleaving. Hence, he is capable of predicting the system utilization and throughput capacity.

Table 3.1: Existing analytical models – Literature review on AVS/RSSs

Publication	System*	Type of vehicle**	Picking	Service time domain***	Arrival process****	Model*****	Performance measures
Malmborg (2002)	AVS/RS	tier-to-tier SL	no	n.a.	n.a.	state equation	avg. values
Malmborg (2003)	AVS/RS	tier-to-tier SL	no	n.a.	n.a.	state equation	avg. values
Fukunari and Malmborg (2008)	AVS/RS	tier-to-tier SL	no	continuous	Poisson	nested queueing model	avg. values
Fukunari and Malmborg (2009)	AVS/RS	tier-to-tier SL	no	continuous	n.a.	CQN	avg. values
Zhang et al. (2009)	AVS/RS	tier-to-tier SL	no	continuous	general	$G G m$ queue	avg. values
Kuo et al. (2007)	AVS/RS	tier-to-tier SL	no	continuous	Poisson	nested queueing model	avg. values
Kuo et al. (2008)	AVS/RS	tier-to-tier SL	no	continuous	Poisson	nested CQN	avg. values
Cai et al. (2014)	AVS/RS	tier-to-tier SL	no	continuous	general	SOQN	avg. values
Ekren et al. (2013)	AVS/RS	tier-to-tier SL	no	continuous	Poisson	SOQN	avg. values
Ekren et al. (2014)	AVS/RS	tier-to-tier SL	no	continuous	Poisson	SOQN	avg. values
Roy et al. (2012)	tier of AVS/RS	tier-captive SL	no	continuous	Poisson	SOQN	avg. values
Roy et al. (2015a)	tier of AVS/RS	tier-captive SL	no	continuous	Poisson	SOQN	avg. values
Roy et al. (2014)	tier of AVS/RS	tier-captive SL	no	continuous	Poisson	SOQN	avg. values
Roy et al. (2015b)	AVS/RS	tier-to-tier SL	no	continuous	Poisson	SOQN	avg. values
Roy and Krishnamurthy (2011)	AVS/RS	tier-captive SL	no	continuous	Poisson	IQM	avg. values
Heragu et al. (2011)	AVS/RS	tier-captive SL	no	continuous	Poisson	OQN	avg. values
Tappia et al. (2017)	SBCSS	tier-captive SL	no	continuous	Poisson	IQM	avg. values

\*AVS/RS: autonomous vehicle storage and retrieval system; SBCSS: shuttle-based compact storage system

\*\*SL: single-level shuttle; ML: multi-level shuttle

\*\*\*n.a.: no value due to used model

\*\*\*\*CQN: open queueing network; CQN: closed queueing network; SOQN: semi open queueing network;

IQM: integrated queueing model consisting of SOQNs and additional queueing systems

Due to the inefficiency of the state equation models, Fukunari and Malmborg (2008) propose an approximate cycle time model for tier-to-tier AVS/RSSs with random storage assignment rules and opportunistic interleaving. They model the system as a nested queueing system consisting of a  $G|G|m^1$  model of the lift nested within the  $M|G|m$  model of the vehicles. Using an iterative procedure

---

<sup>1</sup> The notation commonly used to classify a queueing system is the Kendall's notation:

$$A|B|m - \text{queueing discipline}$$

According to Bolch et al. (1998),  $A$  and  $B$  indicate the distributions of the inter-arrival and service times, and  $m$  denotes the number of servers. In this context,  $M$  describes distributions that are exponentially distributed (Markov process), and  $G$  distributions that are generally distributed. The queueing discipline determines the strategy for the selection process of the next customer when a server becomes available. If no discipline is given, then it is assumed to be FCFS. Moreover, the notation can be extended in various ways such as the number of places in the queue or the arrival and service in batches.

that results in the state distribution of the number of transactions waiting for a vehicle and in service, they approximate the proportion of dual command cycles. In order to do this, they take a binomial perspective of the transactions queue, i.e. they compute the probabilities that there are storage and retrieval transactions in the transaction queue.

To analyze vehicles that might be in failure mode or in maintenance, and that are able to leave the AVS/RS, Fukunari and Malmborg (2009) model a tier-to-tier AVS/RS as a closed queueing network (CQN). They assume opportunistic interleaving and use mean value analysis to determine the average vehicle and lift utilization. The proportion of dual command cycles they approximate by the probability that there is no vehicle at the idle vehicle node. Due to the CQN approach, they are not able to determine the transaction waiting times.

The impact of a non-Poisson arrival process and non-exponential service times is analyzed by Zhang et al. (2009). They assume that the lift is both directly available when needed and seized during the whole transaction by the vehicle. Hence, they model the system as a  $G|G|m$  queue with  $m$  representing the number of vehicles.

Similar to previous studies, Kuo et al. (2007) develop an efficient conceptualization model to analyze tier-to-tier AVS/RSs with random storage and POSC dwell point rules. However, in comparison to other studies, they assume that the lift is seized by the vehicle during the transaction cycle, which means that the lift waits for the vehicle to return to the lift even though other transactions are also requesting the lift. They model the system as an  $M|G|m$  queue with a nested  $G|G|m$  queue that is used to determine the lift waiting times. Although the nested  $G|G|m$  queue leads to substantial errors in estimating the lift waiting time, the vehicle utilization is approximated sufficiently accurate for system conceptualization studies.

Kuo et al. (2008) also use an  $M|G|m$  queue to model an AVS/RS with tier-to-tier vehicles. Different than Kuo et al. (2007), they use a closed queueing network model to determine the lift waiting time of the vehicles. In the closed queueing network, the vehicle nodes correspond to vehicles being either idle or traveling horizontally, and the lift node corresponds to vehicles traveling vertically or waiting for the lift. To solve the closed queueing network, they use mean value analysis. In addition, they model the system such that systems with class-based storage policies can be analyzed. In comparison to simula-

tion results, the model yields accuracy levels adequate for the purpose of the design phase.

Furthermore, there is a rich literature on tier-to-tier configurations, in which the AVS/RS is modeled as a semi-open queueing network (SOQN) to approximate the external queue length more accurately. For example, Cai et al. (2014) use an SOQN approach to determine the average queue length of a tier-to-tier AVS/RS with general service and inter-arrival time distributions. They model the AVS/RS as a multi-class SOQN, which is then aggregated into a single-class SOQN. In analogy to Buitenhok et al. (2000), they suggest a decomposition-aggregation method to reduce the problem to a  $PH|\mu(v)$  queue, where  $PH$  denotes a phase-type distributed arrival process, and  $\mu(v)$  indicates a service rate  $\mu$  that is dependent on the number of customers being served  $v$ . They conclude that the method works well for moderate- and light-load cases. However, the errors obtained by comparison to simulation are greater in the heavy-load case (especially for the number of customers outside the SOQN).

Ekren et al. (2013) present an approximate method to evaluate the performance of an AVS/RS with tier-to-tier vehicles. They aggregate the different customer types of the network into one class and solve the resulting SOQN by using the load-dependent throughput rates of the underlying CQN as service rates of an  $M|M|1$ , birth and death process. The CQN is solved by an extension of Marie's approximation (Ekren and Heragu (2010a)).

Identical to Ekren et al. (2013), Ekren et al. (2014) model the AVS/RS with tier-to-tier vehicles as an SOQN. However, they use a matrix geometric method (MGM) approach to solve the network. First, they reduce the network into a two-station network by using the extended version of Marie's approximation (Ekren and Heragu (2010a)). Afterward, they approximate the aggregated stations by a single station with load-dependent, exponentially distributed service times, and apply the MGM to solve the SOQN (the other station is approximated by a phase-type distribution). In general, they obtain better results in terms of the accuracy of the values for the external queue length compared to the previous model proposed in Ekren et al. (2013).

Also single tiers of an AVS/RS can be modeled as an SOQN. For example, Roy et al. (2012) analyze different design parameters of a tier of an AVS/RS by modeling the tier as an SOQN. To obtain the performance measures, they use a decomposition approach that yields two sub-systems: a closed product

form queueing network for the case when vehicles wait for a transaction, and a single server queue with deterministic service times for the case when the transactions wait for an idle vehicle. Finally, they investigate different design configurations such as depth/width ratio of the tier, vehicle assignment rules, and number of zones.

Similar to Roy et al. (2012), Roy et al. (2015a) analyze design parameters of a tier of an AVS/RS. By modeling a single tier of an AVS/RS as an SOQN, they analyze the impact of different dwell point strategies and a varying location of the cross-aisle on the performance measures. To obtain the performance measures, they use a similar decomposition approach as Roy et al. (2012). After analyzing different system configurations with multiple aisles and one cross-aisle, they find that a return to the load/unload point policy performs better than a POSC dwell point policy. Moreover, they show for different example configurations that as the distance of the cross-aisle from the front of the rack increases, the cycle time improves marginally till approx. 15% from the rack front and then deteriorates. However, like the previous analytical methods on AVS/RSSs, they do not consider vehicle interference effects.

Roy et al. (2014) are the first to model vehicle interference effects that can occur within a tier of an AVS/RS. As a result, they propose blocking protocols that are used to model vehicle blocking within an aisle, within a cross-aisle, or at the intersection of an aisle and cross-aisle. Based on the protocols, they develop an SOQN model of the tier of an AVS/RS and use a similar decomposition approach as Roy et al. (2012) to determine the performance measures of the network. Given the decomposition approach, they analyze the impact of varying configuration parameters such as the number of storage locations, the number of vehicles, and the depth/width ratio on the blocking delays and the transaction cycle time. They show for different tier configurations that the neglect of the blocking delays leads to a significant underestimation of vehicle utilization and cycle time. As expected, the blocking delays increase with the number of vehicles in the tier.

In an extension of Roy et al. (2014), Roy et al. (2015b) model an AVS/RS with tier-to-tier vehicles and two possible devices for vertical movement, namely a lift and a conveyor system that connects the different tiers. To decrease the complexity of the model, they model the tier subnetworks as load-dependent stations. The load-dependent service time they obtain by solving the closed

queueing network model of a single tier. Afterward, they model the AVS/RS as an SOQN and use a decomposition approach that is similar to the approach of Roy et al. (2014) to determine the performance measures. However, since the CQN that needs to be solved for the decomposition approach, does not have a product form solution, they use the approximate mean value analysis to obtain the conditional performance measures. The method considers the above mentioned blocking protocols and can be used for the performance evaluation of tier-to-tier AVS/RSs with different devices for vertical movement such as lifts or conveyor systems.

Another extension of Roy et al. (2014) is used to determine the performance measures of an AVS/RS under blocking consideration with tier-captive vehicles. Roy and Krishnamurthy (2011) connect the method to determine the performance measures of a single tier (see Roy et al. (2014)) with an  $M|G|1$  lift model to incorporate vehicle blocking effects in multi-tier AVS/RSs. Moreover, they analyze the usage of a conveyor vertical transfer mechanism by replacing the lift model by  $G|G|1$  conveyor models with deterministic service times. They show that the throughput can be substantially improved by choosing a conveyor vertical transfer mechanism.

To analyze the impact of tier-captive vehicles, Heragu et al. (2011) model the AVS/RS as an open queueing network (OQN), in which the transactions are modeled as the customers, and the lifts and vehicles as the queueing systems. They use the manufacturing performance analyzer (MPA, see Meng and Heragu (2004)), which is an extension of the queueing network analyzer (QNA) of Whitt (1983), to compute the utilization and the average queue length of the lift and vehicle servers.

In addition to the literature on AVS/RSs, there is a publication of Tappia et al. (2017) on a system that is very similar to AVS/RSs. Hence, it can also be used to analyze AVS/RSs. In their work, they use an SOQN approach for modeling shuttle-based compact storage systems (SBCSSs). The difference to AVS/RSs is that in shuttle-based compact storage systems the storage positions/lanes are multi-deep and the shuttles are able to drive into the lanes to pick up the goods. While the so-called generic shuttles are able to travel on the cross-aisle, specialized shuttles need to be picked up by a transfer car. The vertical movement can be done either by a continuous lift (similar to a conveyor) or by a discrete lift. A single tier is modeled as an SOQN, which is analyzed by first

aggregating all stations into two stations and then applying the MGM. Multi-tier systems are modeled by multiple semi-open queuing networks representing the tiers and a server representing the lift. To obtain the departure process of transactions from the tier and the lift, they approximate the SOQN by a multi server queue.

### 3.1.2 Simulation studies

Due to the complexity of AVS/RSs, there are also many publications on AVS/RSs using simulation for performance evaluation. In the following, we briefly present the relevant literature.

Fukunari et al. (2004) show in their simulation study that there can be significant improvements for AVS/RSs with tier-to-tier vehicles by combining two dwell point policies (return to I/O and last transaction floor) in a real-time control system. The decision, which policy to use at the current moment, is based on a decision tree approach.

Krishnamurthy et al. (2010) analyze the impact of vehicle interference delays within a single tier of an AVS/RS. Therefore, they introduce blocking protocols (see Roy et al. (2014)) that lead to the avoidance of deadlocks within the tier. For a given configuration of a tier of an AVS/RS, they vary the transaction arrival rate and quantify its impact on the performance measures average cycle time, utilization, average number of transactions waiting, and blocking delays. As a result, blocking delays are responsible for 17% to 21% of the transaction cycle time, indicating that vehicle interference effects should be considered when modeling an AVS/RS.

Ekren and Heragu (2010b) use a simulation-based regression analysis to investigate the relationship between the rack configuration and the performance measures of an AVS/RS with tier-to-tier vehicles given a defined capacity constraint. In the model, they consider the average cycle time of the transactions, the average waiting times, and the average utilization of vehicles and lifts. The input variables are the number of tiers, aisles and storage locations on either side of the aisles. To fit the regression functions, they use stepwise regression and best subsets.

Similar to Ekren and Heragu (2010b), Ekren and Heragu (2011) present a simulation-based performance analysis of an AVS/RS with tier-to-tier vehi-

cles. They analyze on the basis of pre-defined rack configurations the impact of scenarios with different numbers of zones/lifts and vehicles per zone/lift on the performance measures average storage/retrieval (S/R) cycle time, and average utilization of shuttle and lift. As a result, they find near optimum levels of shuttle/lift combinations for each rack scenario by searching for the minimum values of the performance measures with the minimum number of lifts and shuttles.

Ekren and Heragu (2012) conduct a simulation study to compare the performance of tier-to-tier AVS/RSs and aisle-to-aisle crane-based AS/RSs on the basis of average flow time, S/R device utilization, average waiting time in the S/R device queue, average number of jobs waiting in the S/R device queue, and costs. Given the needed throughput and storage capacity, they create different practical design scenarios by varying the rack design and the number of S/R devices. First, they simulate the AVS/RS and define various scenarios to select a near-optimal combination of the number of lifts and vehicles. Given that, they conduct a design of experiment under these pre-defined numbers of lifts and vehicles, and four independent factors. The factors are the dwell point policy, the scheduling rule, the input/output locations, and the interleaving rule. By simulating the system for different arrival rate scenarios, they evaluate the main and interaction effects of the factors using analysis of variance (ANOVA). The analysis is based on the performance measures average cycle time for storage and retrieval transactions, average vehicle utilization, and average lift utilization. They find that systems with a point of lift location dwell point strategy, a shortest destination time scheduling rule, an I/O point that is near the middle aisle, and an opportunistic interleaving rule lead for all three performance measures and all arrival rate scenarios to the best results.

An alternative vehicle technology for AVS/RSs is analyzed by Kumar et al. (2014). Using simulation, they investigate the impact of zone-captive vehicles, i.e. vehicles that are dedicated to a fixed number of tiers. Lifts are only used to transfer the goods between the tiers and empty vehicles to other tiers of their zone. They create a large number of different system configurations under random and class-based storage assignment rule and compute the transaction cycle times. They show that zone-captive systems can combine the best features of tier-to-tier and tier-captive configurations, which means that zone-captive systems can lead to shorter transaction cycle times. They also investigate different aisle orientations. For example, moving the cross-aisle to the middle of the tier

and changing the alignment of aisles along the horizontal direction leads to a better system performance. Moreover, they find that horizontal zoning results in even greater benefits in cycle time reduction.

## 3.2 Publications on SBS/RSS

As in the previous section on AVS/RSSs, we will first present the existing analytical models of SBS/RSSs. The analytical models can be classified into basic travel time models for estimating the limiting throughputs of the lifts and vehicles, and queueing models that cover the dependencies between vehicles and lifts, and additionally consider waiting times in the transaction time computation. An overview of these analytical models is given in table 3.2. Subsequently, we present simulation-based literature for performance evaluation of SBS/RSSs.

### 3.2.1 Analytical models

For a basic SBS/RSS configuration with two lifts per aisle that operate independently of each other, Lerher et al. (2015) present a travel time model that enables the computation of the average travel times for the single and dual command cycles of lifts and vehicles.

Based on the existing method, Lerher (2016c) presents in his publication travel time models for SBS/RSSs with double-deep storage. He assumes that the goods are first stored to the second position, i.e. blocking of goods while retrieval is only possible given a fill rate of more than 50%. If the fill rate is larger, blocked goods are rearranged to the nearest free storage location. On the basis of these system dynamics, he develops travel time models for lifts and vehicles performing both single and dual command cycles. Finally, he analyzes the system performance in terms of limiting throughput for different system configurations. He finds that the system performance of the investigated configurations mainly depends on the lift throughput.

Table 3.2: Existing analytical models – Literature review on SBS/RSs

Publication	System*	Type of vehicle**	Picking	Service time domain***	Arrival process***	Model****	Performance measures
Lerher et al. (2015)	SBS/RS	tier-captive SL	no	n.a.	n.a.	travel time model	avg. values
Lerher (2016c)	SBS/RS	tier-captive SL	no	n.a.	n.a.	travel time model	avg. values
Lerher (2016a)	SBS/RS	tier-captive ML	no	n.a.	n.a.	travel time model	avg. values
Lerher (2016b)	SBS/RS	tier-captive SL	no	n.a.	n.a.	travel time model	avg. values
Borovinsek et al. (2017) VDI-Richtlinie 2692 (2015)	SBS/RS SBS/RS	tier-captive SL, tier-captive SL, tier-to-tier SL	no no	n.a. n.a.	n.a.	MOOM travel time model	avg. values avg. values
Liekenbrock (2016)	SBS/RS	tier-to-tier SL, tier-to-tier ML	no	n.a.	n.a.	travel time model	avg. values
Sari et al. (2014)	SBS/RS	tier-captive SL	no	n.a.	n.a.	travel time model	avg. values
Hu et al. (2005)	SP-AS/RS	tier-captive SL	no	n.a.	n.a.	travel time model	avg. values
Hu et al. (2010)	SP-AS/RS	tier-captive SL	no	n.a.	n.a.	LSM	avg. values
Marchet et al. (2012)	SBS/RS	tier-captive SL	no	continuous	Poisson	OQN	avg. values
Kartnig and Oser (2014)	SBS/RS	tier-captive SL	no	continuous	Poisson	$M G 1$ queue	avg. values
Eder and Kartnig (2016b)	SBS/RS	tier-captive SL	no	continuous	Poisson	$M M 1 K$ queue	avg. values
Eder and Kartnig (2016a)	SBS/RS	tier-captive SL	no	continuous	general	$M G 1 K$ queue	avg. values
Epp et al. (2017)	SBS/RS	tier-captive SL	no	discrete	general	OQN	distributions
Zou et al. (2016)	SBS/RS	tier-captive SL	no	continuous	Poisson	FJQN	avg. values

\*SBS/RS: shuttle-based storage and retrieval system; SP-AS/RS: split-platform automated storage and retrieval system

\*\*SL: single-level shuttle; ML: multi-level shuttle

\*\*\*n.a.: no value due to used model

\*\*\*\*OQN: open queueing network; MOOM: multi-objective optimization model with integrated cycle time model; LSM: load shuffling model; FJQN: fork-join queueing network

Furthermore, Lerher (2016a) develops travel time models for SBS/RSs with multi-level shuttles<sup>2</sup> that can serve more than one level. The proposed model is used in a case study to quantify the difference in performance measures given different shuttle velocity profiles. As expected, short traveling times are achieved with the shuttles having fast drives in the horizontal traveling direction.

In addition, Lerher (2016b) presents in his book an energy model to compute the amount of energy consumption and energy regeneration of SBS/RSs based on the number of tiers, aisles, and columns as well as the velocity profiles of the lifts and shuttles.

<sup>2</sup> Referred to as multi-tier shuttle carriers in the publication

Borovinsek et al. (2017) show in their paper that travel time models of SBS/RSSs can also be used as basis for a multi-objective optimization model for the design of SBS/RSSs. The model aims to minimize the following three objective functions for a warehouse that must have a defined minimum storage capacity: the average throughput time, the total cost, and total energy consumption. The design variables, for which predefined bounds are given, are the number of aisles, tiers and columns as well as the velocity/acceleration of shuttles and lifts. To solve the constrained multi-objective problem, they use the non-dominated sorting Genetic Algorithm II (Deb et al. (2002)). For the computation of the average throughput time of the SBS/RS, they use travel time models to determine the average throughput times of the lifts and vehicles as well as the aisles. However, they do not consider waiting processes and, therefore, they are not able to compute the retrieval transaction time.

Travel time models of tier-captive and tier-to-tier SBS/RSSs are also described in the guideline VDI-Richtlinie 2692 (2015) of the Association of German Engineers (VDI). Moreover, they describe the technical components of SBS/RSSs in detail. The guideline is used by Liekenbrock (2016) to compare AS/RSs with tier-to-tier single-level and multi-level<sup>3</sup> SBS/RSSs based on throughput and energy consumption. They find that multi-level SBS/RSSs are a good alternative to single-level SBS/RSSs due to similar performance measures.

Another publication on travel time models for SBS/RSSs is presented by Sari et al. (2014). In their work, they develop single and dual command cycle time models for SBS/RSSs and conduct a simulation study to validate these models. By using two simulation protocols, they validate different speed profiles of the lifts and shuttles.

There are also two publications on travel time models for a split-platform AS/RS (SP-AS/RS) that can be used to handle extra heavy loads (e.g., containers) since the vertical and horizontal movements are carried out by separate devices. The system dynamics are very similar to the ones of tier-captive SBS/RSSs. Hence, the travel time model of Hu et al. (2005) can also be used to determine the average travel time of SBS/RSSs.

---

<sup>3</sup> Referred to as small SRMs in the publication

Moreover, Hu et al. (2010) develop load shuffling algorithms for SP-AS/RSs. Given a batch of retrieval transactions with known retrieval sequence, their method aims to minimize the retrieval response times. They show that the approach can yield a significant increase in system performance by shuffling the loads within the rack during off-peak hours.

In addition to the literature that is focusing on travel time models, there are publications that model SBS/RSs as queueing systems or queueing networks. As a result, it is possible to compute the waiting times. Marchet et al. (2012) present a study in which they determine the average retrieval transaction time of an SBS/RS. They assume that there are only retrieval transactions in the system, and model the system as an OQN. Subsequently, they use a decomposition approach in the continuous time domain to determine the transaction waiting times.

Kartnig and Oser (2014) compute the limiting throughput of an SBS/RS taking into consideration finite buffers between the lifts and the shuttles. In order to consider possible blocking effects, they model both the input and output process of an aisle as an  $M|G|1$  queue. Hereby, they use the lift and shuttle cycle times to compute the arrival and service rates of the systems. By calculating the waiting times, they approximate the blocking delays of the SBS/RS. Given the cycle and waiting times, they compute the limiting throughput of the SBS/RS.

Eder and Kartnig (2016b) compute the limiting throughput of an SBS/RS with multi-deep storage locations and finite buffers between the lifts and the shuttles. In order to consider possible blocking effects, they model the tier of an SBS/RS as an  $M|M|1|K$  queue<sup>4</sup>, i.e. they approximate the inter-arrival and the service times by exponential distributions and assume customers that cannot enter the queue to be lost. They use the model to quantify the impact of multi-deep storage locations on the throughput and the footprint of the SBS/RS.

For a better approximation of the shuttle service time distribution, Eder and Kartnig (2016a) model the tier of an SBS/RS with single-deep storage locations and finite buffers between the lifts and the shuttles as an  $M|G|1|K$  queue. Compared to the previous study of Eder and Kartnig (2016b), this modeling approach leads to a smaller approximation error.

---

<sup>4</sup> In this context,  $K$  represents the maximum number of customers in the system

The only work that focuses on the computation of the complete retrieval transaction time distribution is the publication of Epp et al. (2017). They use a discrete-time queueing network approach to performance evaluation of SBS/RSSs. First, a single aisle of the system is modeled as a discrete-time OQN. Afterward, a decomposition technique is applied, which divides the network into  $G|G|1$  single server stations. Hence, they are able to use the data obtained from the evaluation of the lift and vehicle movements to model the service times by general discrete distributions. Moreover, their approach allows for the computation of the complete probability distributions of the performance measures such as the retrieval transaction time and the number of transactions waiting to be stored. In terms of the approximation quality, the decomposition approach produces lower average errors in comparison to existing OQN approaches in the continuous time domain.

The effects of a parallel processing policy are analyzed by Zou et al. (2016). Under this policy, retrieval transactions directly request the vehicle and lift resources upon arrival. To compute the performance measures of a system with such a control policy, they model the system as a fork-join queueing network (FJQN). Because of the generally distributed service times, the network is non-product form. Hence, they develop an approximation method based on the decomposition principle of the fork-join network (see Bolch et al. (1998)). By carrying out numerical experiments, they compare parallel and sequential processing policies by comparing their results to the results obtained by the model of Marchet et al. (2012). They show for various system configurations that in small systems the parallel processing policy always outperforms the sequential processing policy in terms of response times of retrieval transactions. However, this reverses in large systems with large transaction arrival rates since the parallel processing policy increases the waiting times for the lift.

### 3.2.2 Simulation studies

Similar to AVS/RSSs, there are also publications that use simulation to analyze SBS/RSSs. In addition to their analytical work, Marchet et al. (2013) develop a design framework for SBS/RSSs which is useful to rapidly identify the most appropriate rack configuration that meets the customer requirements (in terms of storage capacity and throughput capacity) at a minimum cost, given the average

retrieval transaction time and the floor space and height constraints. Furthermore, they apply the simulation approach to a case study. In the case study, they show that the framework always identifies the optimal solution among the potential solutions.

Lerher et al. (2015) analyze the system performance of an SBS/RS for different system configurations (rack design, velocities of shuttles and lifts) by conducting a simulation study. The results of the experiments indicate that the throughput capacity significantly depends on the throughput of the lift.

Ekren et al. (2015) conduct a simulation study to analyze preferable rack configurations of SBS/RSs with class-based storage assignment rules. To lower the lift utilization, they assign items with frequent retrievals at the tiers that are closest to the I/O point. The performance measures of interest are the utilization of lifts and vehicles as well as the cycle times of storage/retrieval transactions. Given a required storage capacity, they create possible rack scenarios and compare the design options based on the performance measures. As a result, they observe that the shuttles, which are storing/retrieving items that are frequently retrieved, tend to become the bottlenecks. Hence, the performance measures are better for systems with a large number of tiers, i.e. a large number of tiers for items that are frequently retrieved.

Recently, Ekren (2017) conducted a simulation study to provide graph-based solutions for the design of an SBS/RS. By varying the number of tiers and aisles as well as the arrival rate, they create different scenarios, for which they compute the average utilization of the lifts and the average cycle time of the S/R transactions. Given this data, they provide graph-based solutions for the design of an SBS/RS, i.e. they visualize the performance measures of the simulation study based on the different design variables.

To analyze SBS/RSs with multiple lifts per aisle, Ning et al. (2016) also use simulation. Instead of one lift in front of the SBS/RS, the system which they model consists of multiple lifts within each aisle. For the lift selection, they use a nearest to the retrieval position policy. In their simulation study, they investigate the impact of different rack alternatives and retrieval rates (no storage transactions were considered) on the transaction cycle time and the overall throughput. Based on the results, they suggest rack configurations that are close to the optimal rack configurations.

### 3.3 Conclusion of the literature review

The literature review shows that there is a rich literature on the performance evaluation of AVS/RSSs and SBS/RSSs. However, it also becomes apparent that the existing analytical models, which can be used during the early design phase of an SBS/RSS, have shortcomings regarding the following points.

First, almost all models are computing only the average values of the performance measures. As a result, they cannot be used to design a system such that the retrieval transaction time is shorter than a given value (e.g., 5 minutes) for a given probability. If the service levels have to be considered during the design phase, knowledge about certain quantiles is required (e.g., 95%). To our best knowledge, the only publication, in which the complete probability distribution is computed, is the publication of Epp et al. (2017).

Second, the modeling of discrete lift and shuttle traveling times, which have upper supports and small coefficients of variation, by distributions in the continuous time domain leads to approximation errors. Examples of distributions in the continuous time domain are exponential and phase-type distributions or general distributions that are described by the first two moments of the distributions. Therefore, most approaches in the literature produce approximation errors due to the assumption of service times in the continuous time domain. Only the publication of Epp et al. (2017) assumes service times in the discrete time domain.

In addition, the development of multi-level shuttles leads to the need for performance evaluation tools that consider multi-level shuttles. The literature, however, only presents travel time models of such systems. Hence, there is a lack of models of SBS/RSSs with multi-level shuttles that can compute the retrieval transaction time distribution.

Finally, the existing methods do not include the picking stations as part of the system. Thus, they do not consider the impact of the picking process, i.e. the re-entrant bins, when describing the arrival process at the SBS/RSS.

After reviewing the literature, it becomes apparent that the approach of Epp et al. (2017) offers the best option to determine the distribution of the performance measures given discrete service times. The methods in the continuous time domain that are able to compute the distributions of the performance measures assume either exponential or phase-type distributed inter-arrival and ser-

vice times, i.e. they have to approximate the service times. However, there is much research work left to analyze systems with different system designs such as separate lifts for storage and retrieval transactions, tier-to-tier vehicles, re-entrant bins, and multi-level shuttles. In addition, an analysis of the impact of the random variables on the design decision is missing as well.

Hence, based on the approach of Epp et al. (2017), we present in this work a general modeling approach for SBS/RSs in the discrete time domain which can be used to analyze different SBS/RS designs. With the approach, we are able to show how to design SBS/RSs given the following requirements: storage capacity, throughput, warehouse dimensions, and retrieval transaction time. In addition, we are able to show the impact of re-entrant bins, multi-level shuttles, and different levels of variability of the random variables on the design decision.

# 4 Decomposition approach

In this chapter, we present the decomposition approach that we use for the performance evaluation of SBS/RSSs. Before we present the approach, we will first provide a brief introduction to basics of discrete-time probability theory (compare Schleyer (2007), Matzka (2011), and Özden (2011)), and explain the underlying queueing models and methods that we use for analyzing SBS/RSSs. For a detailed introduction into queueing theory in the continuous and discrete time domain the reader is referred to Bolch et al. (1998) and Tran-Gia (1996), respectively.

## 4.1 Basics of discrete-time probability theory

In contrast to continuous-time stochastic modeling, discrete-time queueing theory assumes time to be discrete. This means that in systems, which are modeled in the discrete time domain, events such as the beginning and the end of service or the arrival of a customer at a queueing system are only recorded at multiples of the constant time increment  $t_{inc}$ .

Hence, the random variables that influence the system behavior are described by discrete random variables. In general, the probability distribution of a discrete random variable  $X$ , which is called probability mass function (pmf), is given by

$$P(X = i \cdot t_{inc}) = \chi_i \quad \forall i = 0, 1, 2, \dots, i_{max}, \quad (4.1)$$

where  $P(X = i \cdot t_{inc})$  denotes the probability that random variable  $X$  takes value  $i \cdot t_{inc}$  and  $\chi_{i_{max}}$  indicates the value of the finite upper support of  $X$ . Probabilities of values smaller than 0 and greater than the upper support are assumed to be 0.

Out of the pmf, we can derive parameters such as the expected value  $E[X]$ , the second moment  $E[X^2]$ , the variance  $\text{VAR}[X]$ , and the squared coefficient of variation  $cv^2[X]$  of the random variable.

$$E[X] = \sum_{i=0}^{i_{\max}} i \cdot \chi_i \quad (4.2)$$

$$E[X^2] = \sum_{i=0}^{i_{\max}} i^2 \cdot \chi_i \quad (4.3)$$

$$\text{VAR}[X] = E[X^2] - E[X]^2 \quad (4.4)$$

$$cv^2[X] = \frac{\text{VAR}[X]}{E[X]^2} \quad (4.5)$$

The squared coefficient of variation (scv) is a measurement for the variability of a random variable. Low values of the scv indicate a low variability, whereas high values of the scv indicate a high variability, i.e. processes with a low scv of the processing time (close to 0) indicate stable processes and processes with a high scv of the processing time indicate unstable processes. For example, the scv of a Poisson process is equal to 1.

The respective cumulative distribution function (CDF) of the random variable  $X$  is given by

$$P(X \leq i \cdot t_{inc}) = \sum_{l=0}^i \chi_l \quad \forall i = 0, 1, 2, \dots, i_{\max}. \quad (4.6)$$

As mentioned before, SBS/RSs are often dimensioned such that the probability of the retrieval transaction time to be smaller than or equal to a given time span is larger than a predefined value (e.g., 95%). Given the CDF of the retrieval transaction time, we determine this probability by the appropriate quantile. The  $u\%-$ quantile of a discrete random variable  $X$ , that gives the value at which the CDF exceeds  $u$  percent, is denoted by  $X_u$ .

$$X_u \Leftrightarrow P(X \leq X_u) \geq u \wedge P(X \leq X_u - 1 \cdot t_{inc}) < u \quad (4.7)$$

The distribution of the sum  $\Phi$  of two independent non-negative random variables  $X$  and  $Y$  can be computed by the convolution of their distributions. In this context,  $\otimes$  is defined as the convolution operator.

$$\Phi = X \otimes Y \quad (4.8)$$

In the discrete time domain, we compute the convolution as follows, where  $\phi_l$  represents the probability that the sum of  $X$  and  $Y$  is exactly  $l$ .

$$\phi_l = \sum_{i=0}^l \chi_i \cdot y_{l-i} \quad \forall l = 0, 1, \dots, l_{max} \quad (4.9)$$

## 4.2 Models and methods in the discrete time domain

According to Schleyer (2007), there are significant advantages when modeling material flow systems in the discrete time domain compared to continuous-time analysis or simulation regarding the accuracy, the level of detail, and the efficiency.

In contrast to most approaches in the continuous time domain, discrete-time queueing analysis allows the description of stochastic processes by more than the first two moments. For a  $G|G|1$  queueing system, Schleyer (2007) showed in experiments, that also higher moments of the inter-arrival and service time distributions such as the skewness and the kurtosis have an influence on the accuracy. Whereas a standard 2-parameter approximation such as the method of Krämer and Langenbach-Belz (1976) obtains the same approximate results for a system with the same values of the first two moments, but different values of the higher moments, the numerical algorithms of Grassmann and Jain (1989) return exact results within an  $\varepsilon$ -neighborhood for any given distribution. Especially in material flow systems, where the service times of the material handling devices often only assume a few discrete values, the representation of the service time by a (possibly multi-modal) discrete distribution leads to a higher accuracy. The discrete distributions that are needed as input for the analysis

can be obtained either by the evaluation of the system dynamics or directly by the results of an as-is analysis.

Moreover, the available discrete-time methods yield the complete probability distributions of the performance measures. This leads to a high level of detail. Thus, a system designer is able to dimension a material flow system not only based on the first two moments of the performance measures, but also on the basis of important quantiles. For example, an SBS/RS can be dimensioned such that the retrieval of a bin does not take longer than three minutes in 95% of the cases. Additionally, buffer capacities can be designed such that the probability of overstock or blocking is below a desired value.

Regarding the efficiency, the numerical discrete-time algorithms are more time-consuming than closed-form solutions in the continuous time domain. However, they are more efficient than simulation studies, which require a long time period for modeling, validation, and performing experiments (see also Schleyer (2007)). Hence, they are well-suited during the early planning stages of material flow systems, in which the system designer wants to compare a large number of possible system designs based on both the average values and the quantiles of the performance measures within a short period of time.

As a result, numerous models and methods in discrete time domain were developed which can be used to analyze processes that are subject to stochastic influences such as material flow systems (see also Matzka (2011)). A selection of discrete-time models and methods is depicted in figure 4.1. Among others, it includes the methods of Grassmann and Jain (1989), Jain and Grassmann (1988), and Furmans and Zillus (1996) that can be used to exactly determine the distributions of the waiting time, the inter-departure time, and the number of customers at the arrival instant of a  $G|G|1$  queueing system, respectively. Given multi-server queueing systems with generally distributed inter-arrival and service times, the methods of Matzka (2011) yield the exact distributions of the number of customers at the arrival instant and the waiting time. Moreover, she presents an approximate approach to determine the inter-departure time distribution.

Since systems with batch arrival and service processes might be of interest, Schleyer and Furmans (2007) developed a method that results in the waiting time distribution of the  $G|G|1$  queueing system with batch arrivals, where the batch size is described by an independent and identically distributed (i.i.d.) ran-

dom variable  $X$ . The distribution of the inter-departure time and the number of customers at the arrival instant can be determined by the methods presented in Schleyer (2007). In terms of the Kendall notation, the system is denoted as a  $G^X|G|1$  queueing system. Methods for more general queueing systems with batch arrival and service processes were developed by Schleyer (2007) and Özden and Furmans (2010). Schleyer (2007) used a decomposition approach to exactly determine the waiting and inter-departure time distribution of a  $G^X|G^{K,K}|1$  queueing system, where the customers first arrive in batches according to the batch size  $X$ , before they are collected and served in batches of constant size  $K$ . The methods of Özden and Furmans (2010) can be used to determine the performance measures of a  $G^X|G^{L,K}|1$  queueing system, where the customers first arrive in batches according to the batch size  $X$ , before they are collected and served when at least  $L$  customers are waiting in the queue. Given all input parameters, the methods yield the distribution of the number of customers at the departure instant, the waiting time, the inter-departure time, and the departing batch size.

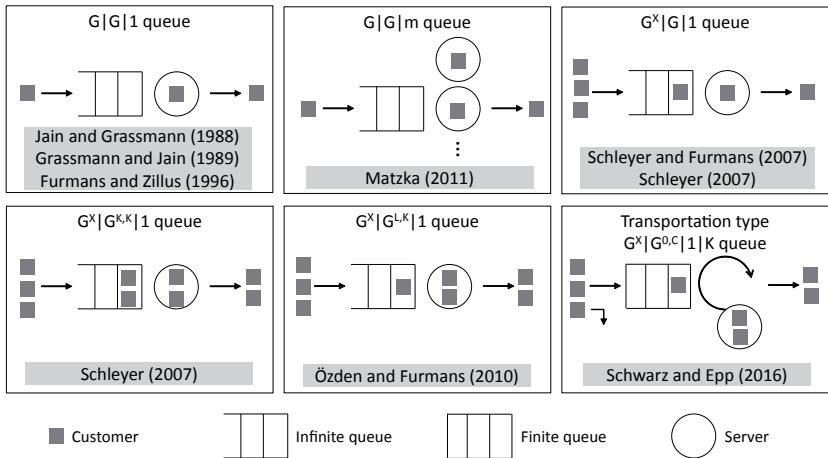


Figure 4.1: A selection of discrete-time models and algorithms

Also bulk queues of transportation type, where the server does not wait for customers, can be analyzed by discrete-time queueing theory. Motivated by the analysis of circulating vertical conveyor systems, Schwarz and Epp (2016) use a Markov chain approach to determine the distribution of the waiting time, the queue length, and the departing batch size of a  $G^X|G^{0,C}|1|K$  queue of transportation type, where customers arrive according to the batch size  $X$  with generally distributed inter-arrival times. They assume a deterministic server capacity  $C$  and a maximum number of customers in the system  $K$ .

The presented models and methods are capable of analyzing systems that can be modeled by a single queueing system in isolation. However, to analyze material flow networks consisting of multiple processing stages, it is necessary that the models take into account the structure of the networks. Hence, models and methods were developed that allow the computation of the network performance measures of open and closed queueing networks in the discrete time domain.

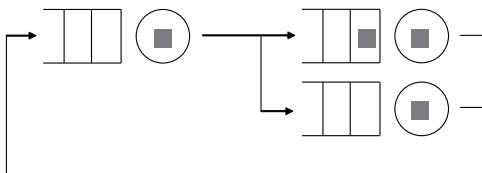


Figure 4.2: An example of a closed queueing network

To analyze closed queueing systems in the discrete time domain, where a fixed number of customers circulates in a closed network consisting of queueing systems with generally distributed service times (see also figure 4.2), Epp et al. (2016) use a discrete-time Markov chain approach to exactly compute the performance measures. They define the system state by the number of customers in the queueing systems and their respective residual service times. Given the discrete service time distributions and the routing matrix of the network, they determine all possible system states and the respective steady-state probabilities. Afterward, they compute for all queueing systems the distribution of the sojourn time and the number of customers. Moreover, they propose an algorithm for the computation of the network cycle time distribution.

Due to the discrete-time Markov chain approach, the state space increases exponentially with an increasing number of customers and queueing systems. Therefore, Epp et al. (2015) developed a decomposition approach to determine the performance measures of closed queueing networks with generally distributed service times. Using this approach, it is possible to approximately determine the performance measures of closed queueing networks with a larger number of customers and queueing systems.

Decomposition approaches are also widely used for the performance analysis of open networks such as the queueing network depicted in figure 4.3.

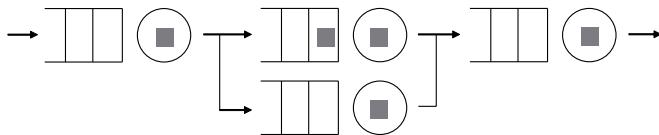


Figure 4.3: An example of an open queueing network

An often used decomposition approach for the analysis of open queueing networks in continuous time domain with generally distributed inter-arrival and service times is the QNA of Whitt (1983). He decomposes the network into  $G|G|m$  queueing systems and computes the performance measures of the single queueing systems as if they were stochastically independent from each other, i.e. he assumes i.i.d. arrivals at the queueing systems. To connect the queueing systems, he uses the departing stream of the upstream systems as arrival stream of the downstream systems. Thus, for every queueing system which is succeeded by another queueing system, he needs to compute the scv of the inter-departure time distribution. To determine the scv of the arrival stream in a network with merge and split operations, he additionally uses methods to compute the scv after the split and merge of stochastic streams. Subsequently, he uses standard methods to compute the performance measures of  $G|G|m$  queueing systems and connects the results to approximately compute the network performance measures. As a result, the QNA is often used to analyze networks with squared coefficients of variation other than 1. In these networks, the errors obtained by the approximate methods of the QNA are smaller than

the errors obtained by using exact methods which assume Markov arrival and service processes.

To analyze open networks with discrete-time queueing models, Furmans (2004) proposes an approach which transfers the idea of Whitt (1983) to the discrete time domain. In analogy to Whitt (1983), Furmans (2004) decomposes the network into stochastically independent  $G|G|1$  queueing systems. But instead of using continuous-time methods, he uses the algorithms of Grassmann and Jain (1989) and Furmans and Zillus (1996) to compute the performance measures of the system, i.e. he assumes discrete inter-arrival and service times. To connect the systems, Furmans (2004) presents models for the stochastic split and merge of customer flows in the discrete time domain. Moreover, he uses the algorithm of Jain and Grassmann (1988) to determine the inter-departure time distribution of a discrete-time  $G|G|1$  queueing system. Hence, the approach enables the system designer to analyze open queueing networks with a higher level of detail than with the QNA.

The level of detail can be further increased by creating a modular building block system that includes the above mentioned discrete-time models of single queueing systems. A first step towards this goal has been achieved by Furmans et al. (2015). In order to analyze large scale material flow systems, they developed a software tool that uses the available discrete-time queueing models and algorithms to determine the sojourn time distributions of these systems. However, they did not create queueing models which are capable of analyzing SBS/RSs. Thus, in this work we develop a decomposition approach that makes it possible to evaluate the performance of SBS/RSs or material handling systems that include SBS/RSs. The approach uses three discrete-time queueing models:  $G|G|1$  queueing system, stochastic split, and stochastic merge. Therefore, we will briefly present these three models in the following subsections. For a detailed description of the underlying algorithms, the reader is referred to Grassmann and Jain (1989), Jain and Grassmann (1988), Furmans and Zillus (1996), and Furmans (2004). Moreover, we develop an approximate method to decrease the computation time of the stochastic split operation. The formal description of this new method, which we denote as the fast split approximation, will be given in section 4.2.4.

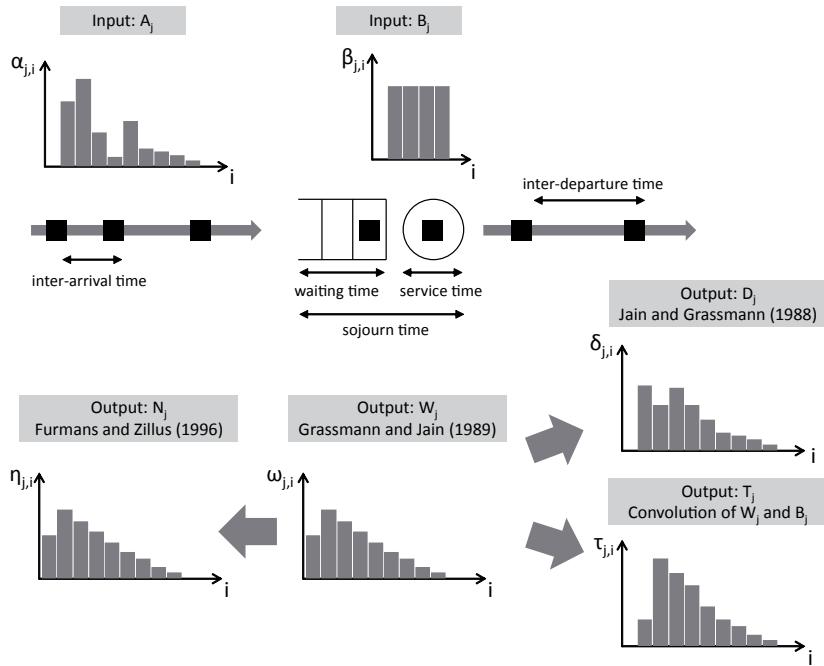
### 4.2.1 $G|G|1$ queueing system

The central element of the decomposition approach is the model of a  $G|G|1$  queueing system in the discrete time domain. It consists of one server and a common queue with unlimited queueing capacity (see figure 4.4). The customers, which belong to a single customer class, arrive individually in batches of one and are served by the single server based on a FCFS discipline. The inter-arrival time between two consecutive customers at an arbitrary queueing system  $j$  is assumed to be i.i.d. according to the discrete random variable  $A_j$ . Moreover, the service time is assumed to be i.i.d. according to the discrete random variable  $B_j$ . Both  $A_j$  and  $B_j$  have finite support and may assume any kind of discrete distribution, including distributions with a multi-modal shape, low scv and discontinuities. The probability for an inter-arrival time of  $i$  time increments is denoted by  $\alpha_{j,i}$ . The probability for a service time of  $i$  time increments is denoted by  $\beta_{j,i}$ .

Hence, the model can be used to evaluate queueing systems with generally distributed discrete inter-arrival and service times. If a customer cannot be served immediately upon arrival, he has to wait in the queue. To determine the waiting time distribution  $W_j$ , Grassmann and Jain (1989) developed an efficient numerical method which is based on the Wiener-Hopf factorization of random walks. In their paper, they present three algorithms for the calculation of the waiting time distribution, which are exact within an  $\epsilon$ -environment. Out of these three algorithms, the third algorithm converges the fastest. Although there is no general proof of convergence for the third algorithm, so far there was no configuration found with a utilization below 1, for which the third algorithm did not converge. Therefore, in the decomposition approach we use the third algorithm of Grassmann and Jain (1989) with  $\epsilon = 10^{-9}$  to compute the waiting time distribution of a  $G|G|1$  queueing system.

Given the distribution of the waiting time, the distribution  $T_j$  of the sojourn time, i.e. the time a customer spends in the queueing system, can be obtained directly by the convolution of the waiting and service time distributions.

Furthermore, Jain and Grassmann (1988) present a method that uses the idle time distribution, which is obtained in addition to the waiting time distribution by the method of Grassmann and Jain (1989), to compute the inter-departure time distribution  $D_j$ . The method is based on the following idea.


 Figure 4.4:  $G|G|1$  model of a queueing system  $j$ 

If a customer arrives at an idle system, the inter-departure time between the customer and his predecessor equals the sum of the idle and service time. Otherwise, if the customer arrives at a busy system, the inter-departure time is equal to the service time of the customer. Therefore, given the idle and service time distributions as well as the probability that an arriving customer encounters an idle system, which equals the probability of a waiting time of 0, it is possible to determine the inter-departure time distribution of the queueing system. In the decomposition approach, the inter-departure time distributions of the upstream queueing systems will be used as input for the determination of the inter-arrival time distributions of the downstream queueing systems.

Given the waiting time distribution, it is also possible to compute the distribution  $N_j$  of the number of customers in the queueing system at the arrival instant of a customer (see Furmans and Zillus (1996)). Hence, we can use the probability distribution of  $N_j$  to identify the needed buffer capacity of a material handling system such that an arriving customer can enter the buffer with a given probability.

### 4.2.2 Stochastic merge

Another central element of the decomposition approach is the model of a merge of two or more than two stochastic streams. It is assumed that the customers of the different stochastic streams arrive in batches of one according to their i.i.d. discrete inter-arrival time distributions. At the merging point, the stochastic streams overlap and form a single stochastic stream. Furthermore, it is assumed that the merging process is done instantly at the merging point without causing a delay. Hence, no queues can arise in front of the merging element and there is the probability of an inter-arrival time of zero time increments of the merged stream if customers arrive at the same time increment at the merging point.

Even though an inter-arrival time of zero time increments might not be feasible for material handling systems, for various reasons it is important to allow inter-arrival times of zero time increments in the decomposition approach. Firstly, if the probability mass for an inter-arrival time of zero time increments is not included in the merged stream and the merged stream is normalized, downstream elements will experience an arrival stream whose throughput is less than the sum of the merged streams' throughput. This will result in a decrease of the utilization of the downstream elements, which results in an underestimation of the throughput time. To avoid this, we could model the stochastic stream as a stream with batch arrivals. However, due to the fact that the computation times of queueing systems that are capable of dealing with batch arrivals are way longer than the computation times needed by the algorithms of Grassmann and Jain (1989), we allow the inter-arrival time distributions to have entries for a probability of zero time increments. In this way, the flow rates in the queueing network will assume the correct expected values and the computation times remain short. Due to the advantages mentioned, and since the other elements of the decomposition approach are able to handle arrival streams with inter-arrival

times of zero time increments, we allow a merged stream with inter-arrival times of zero time increments.

An example of a merge of two stochastic streams is given in figure 4.5. Two stochastic streams arrive with inter-arrival times distributed according to the i.i.d. random variables  $A_\xi$  and  $A_\zeta$ . They merge and form a single stochastic stream which is distributed according to the random variable  $A_j$ . As for the  $G|G|1$  queueing system, the random variables have finite support and may assume any kind of discrete distribution.

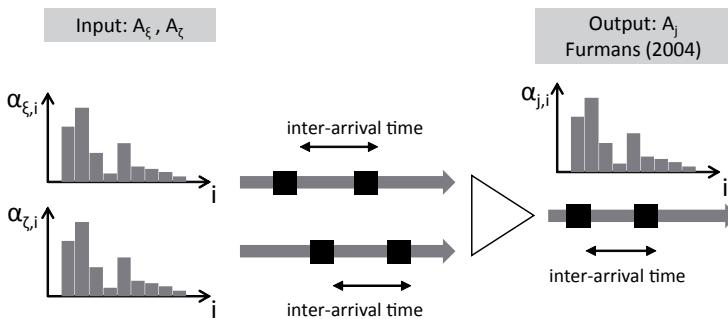


Figure 4.5: Model of a stochastic merge

So far, no method is available that is able to exactly determine the distribution of the merged stream of two or more than two stochastic streams whose inter-arrival times may assume any kind of general discrete distribution. Therefore, in order to determine the inter-arrival time distribution of the merged stream, in the decomposition approach we use an approximation that is based on the assumption that the merged stream is a renewal process. The main steps of the method to merge two stochastic streams, which are described by Furmans (2004) in detail, are presented in the following.

1. Computation of the residual time distribution of the incoming stochastic streams: based on the inter-arrival time distributions, it is possible to determine for an arbitrary time increment the probability that the next customer of an arrival stream arrives in  $0, 1, 2, \dots$  time increments at the merging point.

2. Overlap of the residual time distributions: in this step, the different residual time distributions are merged and an overall residual time distribution of the overlapped stochastic streams is computed. By approximating the merged stream as renewal process, the minimum distribution of the two residual time distributions is computed.
3. Computation of the inter-arrival time distribution of the merged stream: finally, by reversing the steps to determine a residual time distribution given an inter-arrival time distribution, it is possible to recursively determine the inter-arrival time distribution of the merged streams given their overlapped residual time distribution. It is assumed that the resulting inter-arrival time distribution is independent and identically distributed.

To merge more than two stochastic streams, an iterative method is proposed. After the first two streams are merged into one stream, this stream is merged with the third stochastic stream. This procedure is repeated until all incoming streams are merged into one stream.

### 4.2.3 Stochastic split

The model of a stochastic split for discrete distributions, which is used in the decomposition approach, is based on the ordinary Markovian split (see Whitt (1983)). It is assumed that the customers arrive in batches of one according to their i.i.d. discrete inter-arrival time distribution. At the splitting point, the stochastic stream is split into two or more than two stochastic streams, which each are independent and identically distributed. Furthermore, it is assumed that the split is done instantly at the splitting point without causing a delay. Hence, no queues can arise in front of the splitting element.

An example of a split into two stochastic streams is given in figure 4.6. A stochastic stream arrives with an inter-arrival time distributed according to the i.i.d. random variable  $A_j$ . Depending on the routing probabilities  $z_{j,\xi}$  and  $z_{j,\zeta}$ , which represent the probabilities that a customer is routed to directions  $\xi$  and  $\zeta$ , respectively (it is assumed that the routing probabilities are independent of the previously routed customers), the stream is split into two streams which are i.i.d. according to the random variables  $A_\xi$  and  $A_\zeta$ , respectively. As for the  $G|G|1$  queueing system, the random variables have finite support and may assume any kind of discrete distribution.

Given the routing probability  $z_{j,\zeta}$ , the main steps of the method (which are described in detail by Furmans (2004)) to determine for one specific direction  $\zeta$  the inter-arrival time distribution  $A_\zeta$  are presented in the following. In order to determine the inter-arrival time distributions of the other directions after the split, the same method can be applied using the corresponding routing probabilities.

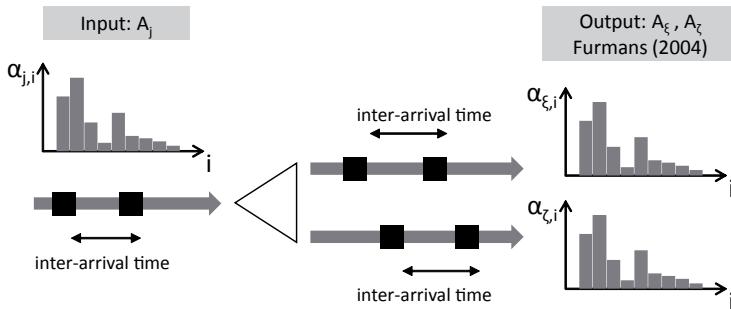


Figure 4.6: Model of a stochastic split

1. Given that the last customer was routed to direction  $\zeta$ , computation of the probabilities  $(1 - z_{j,\zeta})^l \cdot z_{j,\zeta}$  that there are  $l = 0, 1, 2, \dots$  customers routed to other directions than  $\zeta$ , before the next customer is routed to direction  $\zeta$ .
2. Computation of the inter-arrival time distribution  $A_\zeta$ : if there are 0 customers in between two successive customers in direction  $\zeta$ , the inter-arrival time distribution  $A_\zeta$  between these two customers is distributed according to  $A_j$ . If there is 1 customer routed to another direction than  $\zeta$  in between two successive customers in direction  $\zeta$ , the inter-arrival time distribution  $A_\zeta$  between these two successive customers in direction  $\zeta$  is distributed according to the summation of  $A_j$  with itself, which can be computed by the convolution  $A_j \otimes A_j = A_j^{2\otimes}$ . In analogy, if there are  $l$  customers routed to another direction than  $\zeta$  in between two successive customers in direction  $\zeta$ , the inter-arrival time distribution  $A_\zeta$  between these two successive customers in direction  $\zeta$  is computed by the

convolution  $A_j^{(l+1)\otimes}$ . Given the probabilities that there are  $l = 0, 1, 2, \dots$  customers routed to other directions than  $\zeta$  in between two successive customers in direction  $\zeta$ , it is possible to determine the inter-arrival time distribution  $A_\zeta$  by generating a mixture distribution.

$$\begin{aligned} A_\zeta = & (1 - z_{j,\zeta})^0 \cdot z_{j,\zeta} \cdot A_j + \\ & (1 - z_{j,\zeta})^1 \cdot z_{j,\zeta} \cdot A_j^{2\otimes} + \\ & \dots \\ & (1 - z_{j,\zeta})^l \cdot z_{j,\zeta} \cdot A_j^{(l+1)\otimes} + \\ & \dots \end{aligned} \quad (4.10)$$

Due to the assumption of a Markovian split, there is the possibility of an infinite number of customers in between two successive customers for direction  $\zeta$ . Therefore, we stop the summation when the summed up probability mass is larger than or equal to a threshold value  $(1 - \varepsilon_{Split})$ . Afterward, we normalize the distribution. To keep the approximation caused by the normalization small, in the decomposition approach we stop the summation when the missing probability mass is smaller than or equal to  $\varepsilon_{Split} = 10^{-9}$ .

Given  $z_{j,\zeta}$ , we can determine how many summation operations  $n_{sum}$  (and therefore also how many convolution operations  $n_{sum}$ ) we have to perform to reach the necessary threshold value  $(1 - \varepsilon_{Split})$ .

$$\sum_{l=0}^{n_{sum}} (1 - z_{j,\zeta})^l \cdot z_{j,\zeta} \geq 1 - \varepsilon_{Split} \quad (4.11)$$

Due to the formulation of the geometric row, the inequation can be further simplified.

$$1 - (1 - z_{j,\zeta})^{n_{sum}+1} \geq 1 - \varepsilon_{Split} \quad (4.12)$$

To determine the minimum number of operations  $n_{sum}$  needed, we compute

$$\begin{aligned} n_{sum} &= \log_{1-z_{j,\zeta}}(\epsilon_{Split}) - 1 \\ &= \frac{\log_{10}(\epsilon_{Split})}{\log_{10}(1-z_{j,\zeta})} - 1 \end{aligned} \quad (4.13)$$

and round  $n_{sum}$  up to the next integer value. For example, given  $z_{j,\zeta} = 0.5$ , we round  $n_{sum}$  up to 29, which means that we need to perform the summation until the term  $(1-z_{j,\zeta})^{29} \cdot z_{j,\zeta} \cdot A_j^{30 \otimes}$ . Hence, we perform 29 convolutions of  $A_j$  with itself. Given a split into 50 directions with the same routing probabilities in every direction, i.e.  $z_{j,\zeta} = 0.02$ , the number of convolutions increases up to 1025. This results in long computation times for systems with small routing probabilities. Therefore, to speed up the convolution operation within the computation of a splitting element in the decomposition approach, we use the Fast Fourier Transformation (FFT, see also Nussbaumer (1982)).

#### 4.2.4 Fast split approximation

Since the employment of the FFT may also result in long computation times, we present in the following a method that can be used to approximately perform a Markovian split in the discrete time domain. We denote this method as the fast split approximation (FSA).

The idea of the approximation is based on the central limit theorem. It states that the CDF of the sum of  $n$  i.i.d. (discrete) random variables with mean  $\bar{x}$  and variance  $\sigma^2$  each, converges towards a CDF corresponding to a normal distribution with mean  $n \cdot \bar{x}$  and variance  $n \cdot \sigma^2$ . Hence, starting from a sufficiently large number of summations  $n_{FSA}$  (i.e. starting from the  $n_{FSA}$ -th convolution), we can approximate the resulting distributions of the convolutions by normal distributions. Given the mean  $\bar{x}$  and variance  $\sigma^2$  of  $A_j$ , we can re-formulate equation 4.10 as follows, where  $\mathcal{N}^*(\bar{x}, \sigma^2)$  represents a normal distribution with mean  $\bar{x}$  and variance  $\sigma^2$ , which is discretized according to the constant time increment  $t_{inc}$ .

$$\begin{aligned}
A_\zeta = & (1 - z_{j,\zeta})^0 \cdot z_{j,\zeta} \cdot A_j + \\
& (1 - z_{j,\zeta})^1 \cdot z_{j,\zeta} \cdot A_j^{2\otimes} + \\
& \dots \\
& (1 - z_{j,\zeta})^{n_{FSA}-1} \cdot z_{j,\zeta} \cdot A_j^{n_{FSA}\otimes} + \\
& (1 - z_{j,\zeta})^{n_{FSA}} \cdot z_{j,\zeta} \cdot \mathcal{N}^*((n_{FSA}+1) \cdot \bar{x}, (n_{FSA}+1) \cdot \sigma^2) + \\
& (1 - z_{j,\zeta})^{n_{FSA}+1} \cdot z_{j,\zeta} \cdot \mathcal{N}^*((n_{FSA}+2) \cdot \bar{x}, (n_{FSA}+2) \cdot \sigma^2) + \\
& \dots
\end{aligned} \tag{4.14}$$

Since the discretization of the normal distributions is time-consuming as well, we can further speed up the procedure by approximating the weighted sum of the values of the normal distributions by a single Gamma distribution  $\gamma^*(k, \theta)$  with shape parameter  $k$  and scale parameter  $\theta$ , which is discretized according to the constant time increment  $t_{inc}$  and weighted with  $w_{sum} = \sum_{l=n_{FSA}}^{n_{sum}} (1 - z_{j,\zeta})^l \cdot z_{j,\zeta}$ . As before, we can determine how many summation operations  $n_{sum}$  we have to perform to reach a desired threshold value  $(1 - \varepsilon_{Split})$  (e.g., with  $\varepsilon_{Split} = 10^{-9}$ ). Afterward, we normalize the distribution.

$$\begin{aligned}
A_\zeta = & (1 - z_{j,\zeta})^0 \cdot z_{j,\zeta} \cdot A_j + \\
& (1 - z_{j,\zeta})^1 \cdot z_{j,\zeta} \cdot A_j^{2\otimes} + \\
& \dots \\
& (1 - z_{j,\zeta})^{n_{FSA}-1} \cdot z_{j,\zeta} \cdot A_j^{n_{FSA}\otimes} + \\
& w_{sum} \cdot \gamma^*(k, \theta)
\end{aligned} \tag{4.15}$$

The discretized Gamma distribution represents an approximation of the mixture distribution, which results out of the weighted sum of the values of the underlying random variables that are distributed according to the discretized normal distributions. After computing the mean  $\hat{x}$  and variance  $\hat{\sigma}^2$  of the mix-

ture distribution<sup>1</sup>, we can approximate the mixture distribution by a Gamma distribution with the same mean and variance, which is discretized according to the constant time increment  $t_{inc}$ . The mean  $\hat{x}$  and variance  $\hat{\sigma}^2$  we compute as follows.

$$\hat{x} = \sum_{l=n_{FSA}}^{n_{sum}} \frac{(1 - z_{j,\zeta})^l \cdot z_{j,\zeta}}{w_{sum}} \cdot (l + 1) \cdot \bar{x} \quad (4.16)$$

$$\hat{\sigma}^2 = \left( \sum_{l=n_{FSA}}^{n_{sum}} \frac{(1 - z_{j,\zeta})^l \cdot z_{j,\zeta}}{w_{sum}} \cdot [(l + 1) \cdot \bar{x})^2 + (l + 1) \cdot \sigma^2] \right) - \hat{x}^2 \quad (4.17)$$

Out of the mean  $\hat{x}$  and variance  $\hat{\sigma}^2$  of the mixture distribution as well as the knowledge about the mean  $k \cdot \theta$  and variance  $k \cdot \theta^2$  of the discretized Gamma distribution  $\gamma^*(k, \theta)$ , we fit the parameters  $k$  and  $\theta$  of the Gamma distribution.

$$k = \frac{\hat{x}^2}{\hat{\sigma}^2} \quad (4.18)$$

$$\theta = \frac{\hat{\sigma}^2}{\hat{x}} \quad (4.19)$$

### 4.3 General modeling approach

The goal of this work is to develop a general modeling approach that allows the computation of the complete probability distributions of the SBS/RS perfor-

---

<sup>1</sup> The mean  $\hat{x}$  and variance  $\hat{\sigma}^2$  of a mixture distribution of  $n_{mix}$  distributions with weights  $w_l$ , means  $\bar{x}_l$  and variances  $\sigma_l^2$  is computed as follows (Frühwirth-Schnatter (2006)):

$$\begin{aligned} \hat{x} &= \sum_{l=1}^{n_{mix}} w_l \cdot \bar{x}_l \\ \hat{\sigma}^2 &= \left( \sum_{l=1}^{n_{mix}} w_l \cdot (\bar{x}_l^2 + \sigma_l^2) \right) - \hat{x}^2 \end{aligned}$$

mance measures. On the basis of the approach of Epp et al. (2017), we develop an approach for SBS/RSs in the discrete time domain that can be used to analyze a large range of SBS/RS designs. In particular, we integrate the connected picking process and the re-entrance of the bins, as well as the use of multi-level shuttles. In chapter 2, we saw that the incoming storage and retrieval transactions create an open flow of transactions through a network of resources inside and outside of the SBS/RS. Hence, we model SBS/RSs as open queueing networks consisting of single server queues. The customers are representing the storage and retrieval transactions, and the queueing systems are representing the parts of the SBS/RSs that have a population constraint of one due to the use of one secondary resource. Subsequently, we decompose the network into independent  $G|G|1$  queueing systems in the discrete time domain, and model the discrete service time distributions based on the type of transactions that are using the resources, the physical dimensions of the SBS/RS, the characteristics of the resources, and the applied control policies. Finally, we compute the overall network performance measures by the use of an iterative procedure, in which the discrete-time methods presented in sections 4.2.1 ( $G|G|1$  queueing system), 4.2.2 (stochastic merge), 4.2.3 (stochastic split), and 4.2.4 (FSA) are applied.

Theoretically, we could also include the other discrete-time queueing models presented in section 4.2. As a result, we would be able to model a larger range of different SBS/RSs. For example, using  $G|G|m$  queueing systems, we could model parts of the system with a population constraint that is larger than one, e.g., tier-to-tier systems with more than one vehicle per aisle. However, the underlying algorithms would lead to long computation times. This would make it difficult to investigate a large number of system configurations consisting of a large number of those queueing systems in a short period of time. Therefore, we only use the decomposition approach to model SBS/RSs that can be decomposed in  $G|G|1$  queueing systems. The performance measures of  $G|G|1$  queueing systems can be obtained quickly due to the fast algorithms of Grassmann and Jain (1989) and Jain and Grassmann (1988).

In the following subsections, we describe the steps of the modeling approach in more detail, and present the made assumptions. Afterward, we apply the approach to typical tier-captive and tier-to-tier configurations (chapter 5).

### 4.3.1 Steps of the modeling approach

#### Step 1: generation of an open queueing network

In the first step, we model the SBS/RS, the connected picking stations, and the conveying system in between the SBS/RS and the picking stations as an open queueing network consisting of single server queueing systems. The customers of this queueing network are representing the storage and retrieval transactions. Moreover, each queueing system represents a part of the system that has a population constraint of one. It means that only one transaction can be served on a given part of the system at a time due to the fact that there is only one secondary resource designated to it. This part of the system can consist of one or multiple primary resources.

Inside the SBS/RS, the secondary resources of interest are the vehicles and lifts. In the case of a tier-captive configuration, the parts of the system that are served by the vehicles or the lifts are represented by queueing systems since they represent the parts of the system that have a population constraint of one. In the case of a tier-to-tier configuration, both the primary resources and the lift of an aisle are used by the vehicle. Due to the population constraint of one, only one transaction at a time is allowed to enter the part of the system that the vehicle is serving. Since the part of the system that the vehicle is serving includes the lift, we do not have to represent the lift resource by a queueing system. Instead, we only need to model the part of the system that the vehicle is serving by a queueing system. Outside the SBS/RS, the secondary resources of interest are the pickers. If a server is busy upon arrival, the arriving customers are waiting in a single queue. The active conveying systems inside and outside of the SBS/RS are not modeled as queueing systems. Instead, they are represented by the queues of the queueing systems or the edges between the queueing systems.

The routing of the transactions through the network of resources depends on the destination of the storage and retrieval transactions, the control policies, and the existing primary and secondary resources that can be used to transport the goods. As a result, the routing of the transactions determines the routing of the customers through the queueing network. The split and merge of transactions is represented by the split and merge of the stochastic customer streams. Given

the arrival rates from outside and the routing of the customers, the arrival rates at each queueing system can be determined.

### **Step 2: modeling of the service time distributions**

In this step, we model the discrete service time distributions of the queueing systems inside the SBS/RS based on the type of transactions that are using the resources, the physical dimensions of the SBS/RS, the characteristics of the resources, and the applied control policies. The discrete service time distributions of the picking stations are assumed to be known (e.g., by the use of time studies). Therefore, they represent an input for the queueing network model.

For every queueing system inside the SBS/RS, we first describe the possible sequences of events during which a transaction uses the resource. We denote these sequences of events as transaction cycles. In many cases, we can classify the transaction cycles based on the type of transaction, the location of the resource upon request, and the storage/retrieval location of the transaction. In general, the type of transaction can be either a storage or a retrieval transaction, and the location of the resource upon request depends on the storage/retrieval location of the preceding transaction and the dwell point strategy of the resource. The storage/retrieval location of the transaction is the position to unload the goods in case of a storage transaction, and the position to load the goods in case of a retrieval transaction. If needed, more variables can be added to classify the transaction cycle. For example, we can add the source of the storage transactions and the destination of the retrieval transactions to the classification. In case of a storage transaction, the source represents the position to load the goods. In case of a retrieval transaction, the destination represents the position to unload the goods. Since there is often just one source and destination, this variable is omitted in many cases.

After the classification, we compute the probability and the service time for every transaction cycle. The probability of a transaction cycle depends on the storage/retrieval transaction ratio and the applied control policies such as storage assignment and dwell point strategies. The service time of a transaction cycle depends on the times to load and unload the bins, the so-called transfer times, as well as the traveling times. Whereas the (un)loading times are assumed to be constant, the traveling times depend on the traveled dis-

tances as well as the acceleration/deceleration rates and velocities of the resources. The traveling times also include the times to change the aisle in case of aisle-to-aisle vehicles and to change tiers in case of tier-to-tier vehicles. In addition, we round the service time to the next time increment due to the discrete-time approach.

Finally, we determine the probability that the service time distribution assumes a value of  $i$  time increments by summing up the probabilities of all transaction cycles that lead to a service time of  $i$  time increments. If we are interested in the service time distribution of the retrieval transactions, we only sum up the probabilities of the retrieval transaction cycles that lead to a service time of  $i$  time increments, and afterward normalize the distribution given the probability that the transaction is a retrieval transaction.

Since many resources have the same characteristics such as transfer times, velocities, and acceleration/deceleration rates, and additionally have to handle the same flow of storage and retrieval transactions due to the physical system design and applied control policies, the service time distributions of many resources tend to be identical. For example, in many systems, the service time distributions of the vehicle queueing systems are identical.

### **Step 3: decomposition of the network into $G|G|1$ queueing systems and computation of the network performance measures**

The performance measures of interest are the utilization of the secondary resources, the retrieval transaction time distribution, the number of transactions waiting to be stored/picked, and the inter-departure time distribution of leaving retrieval transactions. Given the customer flow of the queueing network, which results in the arrival rates at the queueing systems, and the service time distributions of the queueing systems, we can directly obtain the utilization of the queueing systems, which represent the respective secondary resources of interest.

To determine the other performance measures of the queueing systems, we decompose the network into independent  $G|G|1$  queueing systems in the discrete time domain. As shown in section 4.2.1, given the inter-arrival and service time distributions of a queueing system, we can determine both the waiting and the inter-departure time distributions using the algorithms of Grassmann and Jain

(1989) and Jain and Grassmann (1988), respectively. Moreover, we are able to compute the sojourn time distribution by the convolution of the waiting and service time distributions, and the number of customers in the queueing system at the arrival instant of a customer by the algorithm of Furmans and Zillus (1996). Subsequently, we can directly derive the number of waiting customers at the arrival instant of a customer. If there are only retrieval transactions entering the queueing system, the sojourn time distribution of the queueing system equals the sojourn time distribution of the retrieval transactions at the queueing system. If there are both storage and retrieval transactions using the queueing system, we determine the sojourn time distribution of the retrieval transactions at the queueing system by the convolution of the waiting time distribution and the service time distribution of the retrieval transactions. If there are only storage transactions using the queueing system, the number of waiting customers at the arrival instant of a customer equals the number of bins waiting in front of the resource at the arrival instant of a bin. Moreover, we use the Binomial distribution if we want to determine only the number of waiting storage or retrieval transactions at the arrival instant of a transaction given the distribution of the number of waiting storage and retrieval transactions.

In a linear topology, we use the inter-departure time distribution of the upstream queueing system as the inter-arrival time distribution of the downstream queueing system. To determine the inter-arrival time distribution of a downstream queueing system given a non-linear topology, we have to split and/or merge the customer streams of the upstream queueing systems. In our decomposition approach, we use the methods described in section 4.2 to perform the split and merge operations.

In topologies, in which the customer stream generates loops or cycles, we determine the steady-state performance measures using an iterative procedure (as it has been done in previous publications on discrete-time queueing networks with loops or cycles, e.g., by Furmans et al. (2015)). In the first iteration, we do not consider the customer streams that are coming from a loop or a cycle in the computation of the inter-arrival time distribution of a queueing system. In the subsequent iterations, however, we consider the customer streams that are coming from a loop or a cycle, and which were determined in the last iteration. For example, the non-empty bins after picking are generating a re-entering stream of storage transactions. In the iterative algorithm, we use the computed distribution of the re-entering storage transaction stream of the last iteration as input

for the computation of the distribution of the storage transaction stream entering the SBS/RS in the subsequent iteration. After every iteration, we compute the retrieval transaction time distribution. If the absolute difference between the expected values of the retrieval transaction time distributions of two successive iterations is bigger than a given threshold value, we continue with the subsequent iteration. Otherwise, we stop the iterations and set the performance measures that are obtained in the last iteration as the steady-state performance measures of the queueing network.

### 4.3.2 Assumptions

The basic building blocks of the decomposition approach are discrete-time  $G|G|1$  queueing systems and the discrete-time methods to split and merge stochastic streams. This leads to the following assumptions and SBS/RS constraints.

- Single customer class: due to the used algorithms, the storage and retrieval transactions form a single customer class. Hence, the customers are routed based on the same routing probabilities and served based on the same service time distributions.
- Infinite queueing capacities: since we use discrete-time  $G|G|1$  queueing systems with infinite queueing capacities to model the resources, we approximate the finite physical buffer places by infinite queues, i.e. we assume that there is no blocking. This assumption leads to small approximation errors for systems, in which the designer chooses the buffer capacities such that the possibility of blocking is close to zero. If the needed buffer capacities are too large, the system designer may choose a system in which the utilization of the resource is lower, thus the queue length is shorter.
- FCFS service discipline: at the queueing systems and the elements to split and merge the customer streams, the customers are served based on a FCFS discipline. Hence, we only model SBS/RSs that are controlled by a FCFS sequencing strategy. Moreover, in case of picking orders with multiple orderlines that do not arrive in the right sequence, we assume that there is enough space at the picking stations to buffer carton boxes and to access the carton boxes in the right sequence.

- Common queue: if a server is busy upon arrival, the arriving customers are waiting in a single queue. Thus, retrieval and storage transactions that might wait virtually or physically in different buffer places, are waiting in the same queue of the queueing system. As stated before, they are served based on a FCFS discipline.
- Batch size of one: the customers are arriving at the queueing systems in batches of one. Furthermore, they are served in batches of one. Due to the FCFS service discipline, the common queue, and the batch size of one, we only model SBS/RSSs that are operated under single command cycles.
- I.i.d. arrival and service processes: the inter-arrival time distributions at the queueing systems and the elements to split and merge the customer streams are assumed to be i.i.d. Also the service time distributions at the queueing systems are assumed to be i.i.d.
- Discrete inter-arrival and service time distributions: both the inter-arrival and service time distributions of the queueing network are assumed to have a finite upper support. They are discretized based on the constant time increment  $t_{inc}$ . Distributions with infinite upper support are truncated and normalized.
- Stochastic routing: the method to split stochastic streams assumes a Markovian split. Moreover, the stochastic merge is based on the assumption that the merged stream is a renewal process.
- No delay at the splitting and merging points: due to the used algorithms to split and merge stochastic streams, we assume that there are no delays and therefore no queues in front of the splitting and merging points.
- Random storage assignment: regarding the storage assignment rule, we assume a random storage assignment.
- Equal access frequency: we assume that the retrieval transactions are equally split among the storage locations. Thus, they induce a leveled workload on the resources. The same applies to the access frequency of the picking stations in case of the retrieval transactions that are routed to the picking stations.
- Single-deep and equally sized storage locations: we assume that the storage locations are single-deep and equally sized.

- No parts with a population constraint larger than one: due to the modeling of the system as an open queueing network consisting of single server queueing systems, we assume that there are no parts of the system with a population constraint larger than one. It means that there can be no more than one shuttle assigned to a specific part of the SBS/RS, and no more than one lift assigned to an elevator shaft.
- Shuttle capacity of one: due to the population constraint of one, we assume that the shuttles have a capacity of one bin. Therefore, we assume that there is only one LHD per shuttle.
- Use of lift for vertical movement with capacity of one: we assume that lifts are used for the vertical movement. Each lift is assigned to one elevator shaft and has one platform that can hold one bin.
- Single-level or multi-level shuttles: we assume that either single-level or multi-level shuttles can be installed.
- Use of active conveying system outside of SBS/RS: we assume that there are active conveying systems used outside the SBS/RS to transport the goods to the picking stations and back to the SBS/RS. Since they are represented by the queues of the queueing systems or the edges between the queueing systems, we do not model the time of the goods on the conveying system. However, because we are only interested in the inter-arrival time distribution of the re-entrant bins, which we also obtain by the given modeling approach, we do not need to model the time of the goods on the conveying system. Furthermore, we assume that the picked items leave the system after the picking process via another material handling system that is of no interest in the scope of this thesis.
- Dedicated resource assignment at the queueing systems and shortest path routing: due to the population constraint of one and the modeling of the parts of the system with a population constraint of one as single server queueing systems, there is no set of resources on a given part of the system that a transaction can choose from after it was routed to this part of the system. This leads to a dedicated resource assignment at this part of the system. Moreover, we assume a shortest path routing strategy.
- POSC dwell point strategy: for the vehicles and lifts, we assume a POSC dwell point strategy.

# 5 Modeling of tier-captive and tier-to-tier configurations

In this chapter, we apply the modeling approach to calculate the performance measures of two system configurations. At first, we model a tier-captive configuration in section 5.1. Afterward, we show how to model a tier-to-tier configuration (see section 5.2).

## 5.1 Tier-captive configuration

The process of the calculation of the performance measures is presented in the same order as explained in chapter 4. At first, we give a brief overview of the physical design, the control policies and the transaction flow of the system under investigation. Afterward, we determine the corresponding queueing model. In section 5.1.3, we show how to determine the service time distributions of the secondary resources. Finally, we present the algorithm to compute the performance measures.

### 5.1.1 System description

The system under investigation is a tier-captive SBS/RS. An example of such a system with three aisles and two picking stations is depicted in figure 5.1. In general, it consists of  $n_a$  aisles,  $n_t$  tiers per aisle,  $n_l$  levels per tier and  $n_c$  storage columns on either side of the aisle. As shown in figures 5.2 and 5.3, the storage columns of any level are equally sized and can hold one bin. The distances between two tiers, two levels of a tier and two storage columns are given by  $d_t$ ,  $d_l$  and  $d_c$ , respectively. Since the shuttles can neither change the aisle nor the tier, the system does not have any cross-aisles.

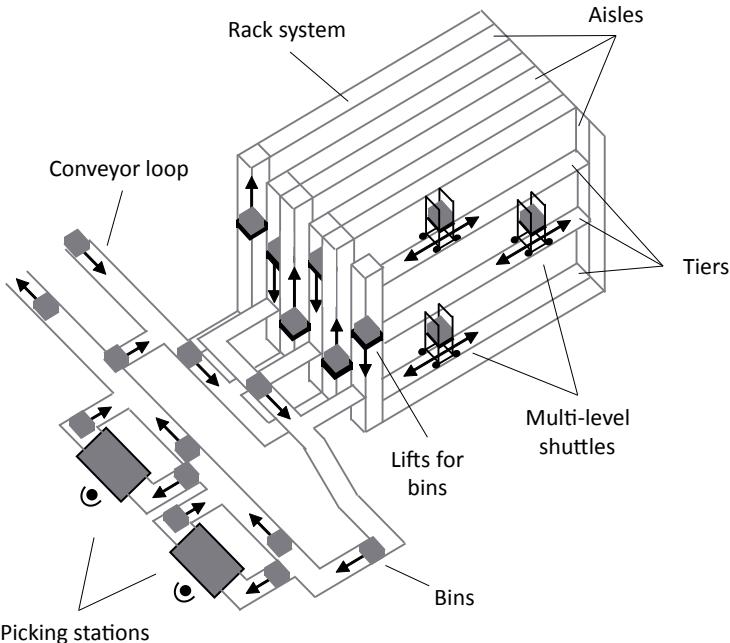


Figure 5.1: Tier-captive system with 3 aisles and 2 picking stations

The shuttles have one load handling device with a capacity of one bin, and can serve the  $n_l$  levels of a tier. They travel simultaneously along the aisle and lift/lower the load handling device in vertical direction towards the target level. Their maximum velocities in horizontal direction along the aisle and vertical direction to lift/lower the load handling device are  $v_{v,x}$  and  $v_{v,y}$ , respectively. The acceleration rates, which are identical to the deceleration rates in the respective direction, are given by  $a_{v,x}$  and  $a_{v,y}$ . The vehicle transfer times  $t_{v,trans}$  to load/unload the bin between the (un)loading points and the shuttle are identical to the transfer times to load/unload the bin between the storage rack and the shuttle.

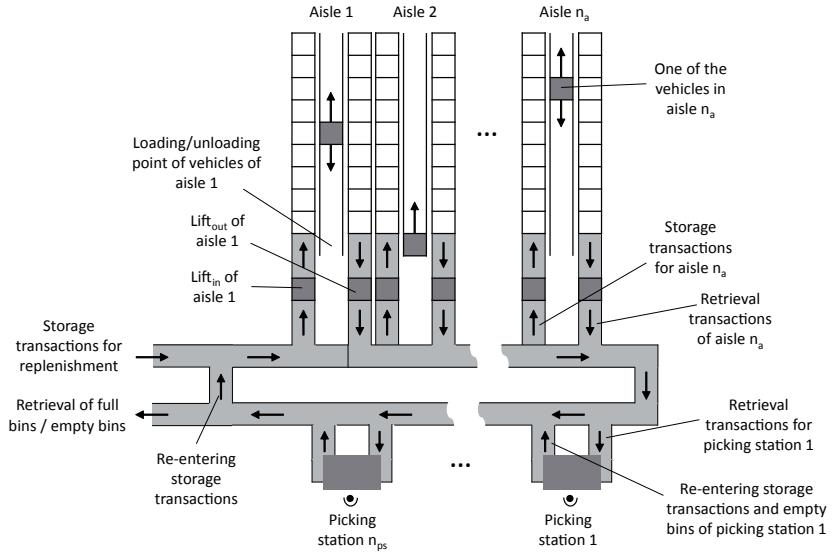


Figure 5.2: Plan view of the tier-captive system

The (un)loading points of the vehicles and the buffers of the incoming and outgoing bins of a tier are located in front of the aisle. In vertical direction, the distance between the (un)loading points and the lowest level of a tier is 0. In horizontal direction, the distance between the (un)loading points and the first column of a tier is equal to  $d_c$ .

In front of that, one incoming lift ( $lift_{in}$ ) provides the vertical transport of the incoming bins from the input point of the aisle to the buffers of the target tiers, and one outgoing lift ( $lift_{out}$ ) provides the vertical transport of the outgoing bins from the buffers of the tiers to the output point of the aisle. Each lift can hold one bin and is only able to travel in vertical direction. Their maximum velocities in vertical direction are  $v_{l_{in}}$  and  $v_{l_{out}}$ , respectively. The acceleration rates, which are identical to the deceleration rates of the respective lift, are given by  $a_{l_{in}}$  and  $a_{l_{out}}$ . The incoming lift transfer times  $t_{l_{in},trans}$  to load/unload the bin between the input point of the aisle and the incoming lift are identical

to the transfer times to load/unload the bin between the incoming lift and the buffers. In analogy, the outgoing lift transfer times  $t_{l_{out},trans}$  are identical for the transfer of the bin between the buffers and the outgoing lift as well as the outgoing lift and the output point of the aisle. The distances between the lowest tier of the system and the input and output points of the aisles are given by  $d_{l_{in}}$  and  $d_{l_{out}}$ , respectively.

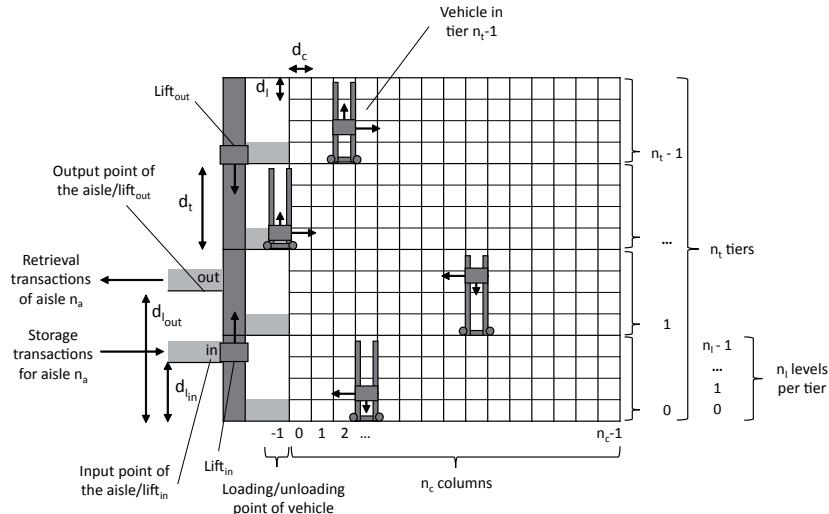


Figure 5.3: Aisle view of the tier-captive system

Outside of the SBS/RS, a conveyor system transports the outgoing bins from the outgoing lifts to their target destination such as one of the  $n_{ps}$  picking stations or another subsequent process. The probability that a bin is routed to one of the picking stations is denoted by  $p_{ps}$ . The same conveyor system transports the empty bins from the picking station to another process in the warehouse that refills the bins. The probability that a bin is emptied at the picking station is denoted by  $p_{em}$ . The conveyor system also transports the re-entering bins from the picking stations and the replenishment bins from the refilling process to the incoming lifts. At each picking station there is one worker that picks the required number of items out of the bin.

The control strategies of the system are as follows. Both the shuttles and the lifts operate independently from each other in a single command cycle mode. They process the arriving storage and retrieval transactions based on a first-come-first-serve strategy. Furthermore, the incoming storage transactions are assigned randomly among the storage locations. The access frequency of the retrieval transactions is equally shared among the storage locations. Also the probability to route a bin to one of the picking stations is equally and independently shared among the picking stations. As a result, the system load is – on average – equally shared between the shuttles, lifts and picking stations. If all waiting transactions are served, both the shuttles and the lifts idle at the point of service completion.

There are two types of transactions, namely retrieval and storage transactions. The retrieval transactions can both represent a transport order for a bin that needs to be transported to a picking station and a bin that needs to be directly transported to another subsequent process. Additionally, there are retrieval transactions for bins that got empty at the picking process and need to leave the system to be refilled. In contrast to that, the storage transactions both represent the transport orders for the re-entering bins after the picking process and the replenishment bins coming from the refilling process.

In the following, we describe the flow of the transactions through the system based on the graph model depicted in figure 5.4. The nodes represent the primary resources inside the SBS/RS as well as the picking stations outside the SBS/RS. The edges represent the flow of transactions between the primary resources inside the SBS/RS. Since we solely focus on the performance measures of the SBS/RS, we do not represent the conveyor system outside the SBS/RS by nodes. Instead, edges represent the conveyor system outside the SBS/RS. The lifts, the shuttles and the pickers are the secondary resources that lead to a population constraint at the primary resources. In figure 5.4, dotted rectangular shapes around the primary resources depict the secondary resources.

## 5 Modeling of tier-captive and tier-to-tier configurations

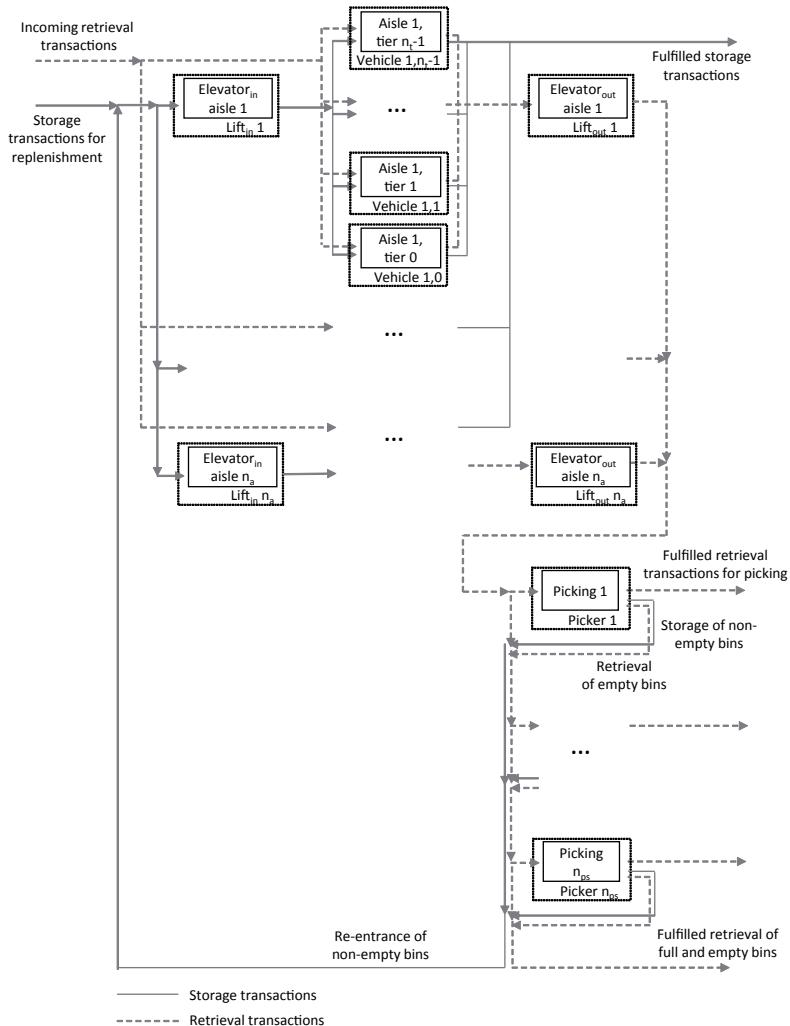


Figure 5.4: Graph model of the tier-captive system

The dashed edges of the graph model represent the flow of retrieval transactions, whereas the solid edges represent the flow of storage transactions. An incoming retrieval transaction first uses a shuttle to transport the requested bin to the buffer in front of the outgoing lift. Afterward, it uses the outgoing lift to transport the bin out of the SBS/RS to the connected conveyor system. If the bin needs to be directly transported to a subsequent process, it uses the conveyor system to leave the system, bypassing all picking stations. If the bin needs to be transported to a picking station, the system routes the bin using the shortest path on the conveyor system to the picking station that requested the bin. At the picking station, the retrieval transaction uses the picker to take the requested number of items out of the bin. At that point, the retrieval transaction for picking is fulfilled and leaves the system.

After the picking process, the system generates a new retrieval transaction if the bin is empty to remove the bin from the system. It uses the conveyor system to leave the system, bypassing all picking stations. If the bin is not empty, the system generates a new storage transaction with a target storage location that is defined by the random storage assignment rule. It uses the shortest path on the conveyor system to transport the bin to the buffer in front of the incoming lift of the target aisle. Also the storage transactions, that enter the system for replenishment, use the conveyor system to route the bins to the buffer in front of the incoming lift of the target aisle. At the elevator shaft, both types of storage transactions use the lift to transport the bins to the incoming buffers at the target tiers. From there, the transactions use the shuttle to store the bin in the rack system. After that, the storage transactions are fulfilled and virtually leave the system.

Table 5.1: Notation used in the system description

$a_{l_{in}}$	Lift <sub>in</sub> acceleration/deceleration rate in vertical direction
$a_{l_{out}}$	Lift <sub>out</sub> acceleration/deceleration rate in vertical direction
$a_{v,x}$	Vehicle acceleration/deceleration rate in horizontal direction
$a_{v,y}$	Vehicle/LHD acceleration/deceleration rate in vertical direction
$n_a$	Number of aisles
$n_c$	Number of columns on either side of an aisle
$n_l$	Number of levels per tier
$n_t$	Number of tiers per aisle
$n_{ps}$	Number of picking stations
$p_{em}$	Probability that a bin is empty after picking
$p_{ps}$	Probability that the target of the retrieval transaction is a picking station
$t_{l_{in},trans}$	Lift <sub>in</sub> transfer time, i.e. time to load or unload the goods onto or from the incoming lift
$t_{l_{out},trans}$	Lift <sub>out</sub> transfer time, i.e. time to load or unload the goods onto or from the outgoing lift
$t_{v,trans}$	Vehicle transfer time, i.e. time to load or unload the goods onto or from the vehicle
$v_{l_{in}}$	Lift <sub>in</sub> velocity in vertical direction
$v_{l_{out}}$	Lift <sub>out</sub> velocity in vertical direction
$v_{v,x}$	Vehicle velocity in horizontal direction
$v_{v,y}$	Vehicle/LHD velocity in vertical direction
$d_c$	Distance between two storage columns
$d_l$	Distance between two levels of a tier
$d_t$	Distance between two tiers
$d_{l_{in}}$	Distance between the lowest tier and the input point of the lift <sub>in</sub>
$d_{l_{out}}$	Distance between the lowest tier and the output point of the lift <sub>out</sub>

---

### 5.1.2 Queueing network model

This section presents the discrete-time queueing network that is used to model the tier-captive SBS/RS (see figure 5.5). In this network, those parts of the system with a population constraint of one are represented by single server queueing systems with unlimited queueing capacities. Therefore, the parts of the system that are served by the vehicles, incoming and outgoing lifts, and the pickers form the set of queueing systems. Furthermore, we assume that the transactions form one customer class and that the stochastic routing of the customers is determined by the routing matrix  $Z$ . The probability of a customer to be routed from queueing system  $j$  to queueing system  $\xi$  is denoted as  $z_{j,\xi}$ . The inter-arrival times of the incoming retrieval transactions are assumed to be independent and identically distributed according to the random variable  $A_R$ . Therefore, the arrival rate of the incoming retrieval transactions is determined as follows.

$$\lambda_R = \frac{1}{E[A_R]} \quad (5.1)$$

Due to the equal access frequency of the retrieval transactions to the storage locations, the routing probability of the retrieval transaction stream to one of the vehicle stations is equal to  $\frac{1}{n_a \cdot n_t}$ . The split of the retrieval transaction stream into  $n_a \cdot n_t$  identical streams leads to the inter-arrival time distribution  $A_{veh,R}$  of the retrieval transactions arriving at each vehicle station.

The inter-arrival times of the storage transactions for replenishment are assumed to be independent and identically distributed according to the random variable  $A_{S_{rep}}$ . Thus, the arrival rate is computed as follows.

$$\lambda_{S_{rep}} = \frac{1}{E[A_{S_{rep}}]} \quad (5.2)$$

Together with the re-entering storage transactions, which are described by the inter-arrival time distribution  $A_{S_{re-en}}$  and the arrival rate  $\lambda_{S_{re-en}} = \frac{1}{E[A_{S_{re-en}}]}$ , they form the stream of the incoming storage transactions. The inter-arrival time distribution of this stream is denoted by  $A_S$ . Its arrival rate is determined as follows.

$$\lambda_S = \lambda_{S_{rep}} + \lambda_{S_{re-en}} \quad (5.3)$$

Due to the random storage assignment rule, the routing probability of the storage transaction stream to one of the incoming lift stations is equal to  $\frac{1}{n_a}$ . The split of the storage transaction stream into  $n_a$  identical streams leads to the inter-arrival time distribution  $A_{lift_{in}}$  of the storage transactions arriving at each incoming lift station. Due to the identical physical design and control policies for all incoming lifts, the service time distributions  $B_{lift_{in}}$ , which we assume to be i.i.d., are identical for each incoming lift station. Therefore, also the inter-departure time distributions  $D_{lift_{in}}$  are identical for each incoming lift station.

After the incoming lift station, the storage transactions arrive at the vehicle stations. Due to the random storage assignment rule, the routing probability to one of the vehicle stations is equal to  $\frac{1}{n_t}$ . The split into  $n_t$  identical streams leads to the inter-arrival time distribution  $A_{veh,S}$  of the storage transactions arriving at each vehicle station. Together with the retrieval transactions, they form the stream of arriving transactions at a vehicle station. The inter-arrival time distribution of this stream, which is identical for each vehicle station, is denoted by  $A_{veh}$ .

Due to the identical physical design and control policies for all vehicle stations, the service time distributions  $B_{veh}$ , which are assumed to be i.i.d., and the inter-departure time distributions  $D_{veh}$  are identical for all vehicle stations as well. Only the retrieval transaction streams are routed to the outgoing lift stations. Therefore, after each vehicle station the retrieval transactions are split from the total departure stream. Due to the equal distribution of incoming storage and retrieval transactions among the vehicle stations, the ratio of retrieval transactions among all departing transactions of a vehicle station is equal to the ratio  $p_R$  of the retrieval transactions to the sum of storage and retrieval transactions, which is determined as follows.

$$p_R = \frac{\lambda_R}{\lambda_S + \lambda_R} \quad (5.4)$$

This ratio is equal to the routing probability of the departing retrieval transaction stream of a vehicle station to the outgoing lift station of the respective aisle. The other part of the departure stream of a vehicle station represents the fulfilled storage transactions that virtually leave the system after the storage of the bins in the rack.

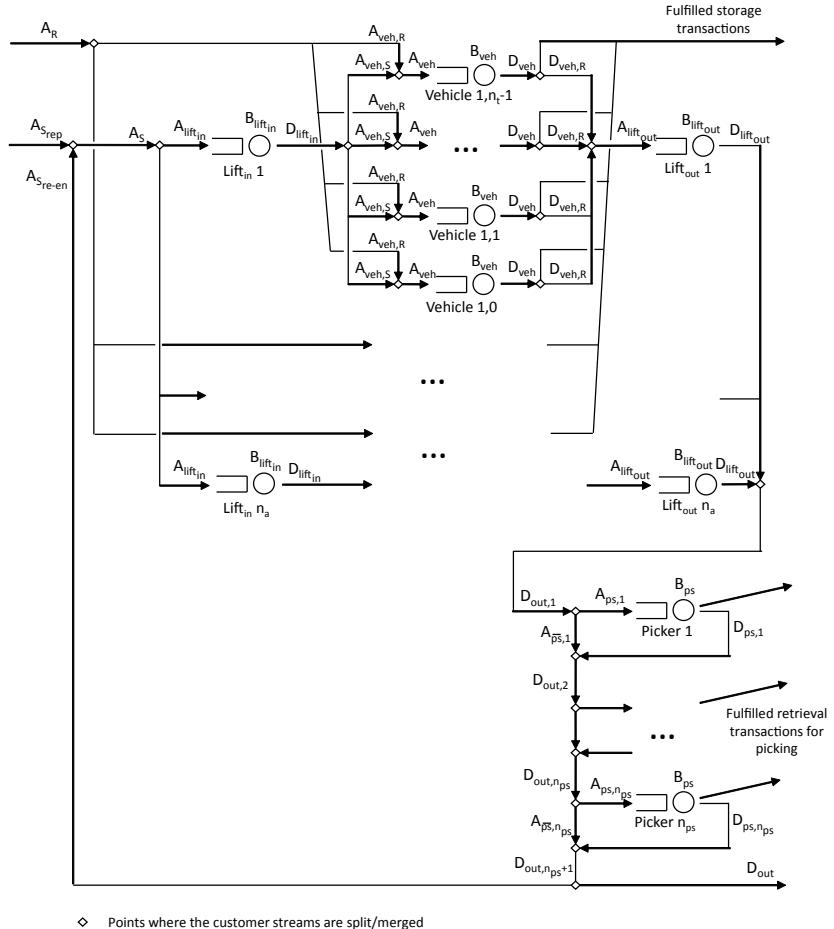


Figure 5.5: Queueing model of the tier-captive system

Following the vehicle stations, the retrieval transactions assigned to an aisle with identical inter-departure time distributions  $D_{veh,R}$  generate an arrival stream of the respective outgoing lift station, which is identical for all outgo-

ing lift stations. The inter-arrival time distribution of an outgoing lift station is denoted by  $A_{lift_{out}}$ . Since the physical design and the control policies are identical for all outgoing lift stations, the service time distributions  $B_{lift_{out}}$ , which are assumed to be i.i.d., and the inter-departure time distributions  $D_{lift_{out}}$  are identical for all outgoing lift stations as well. All inter-departure time distributions  $D_{lift_{out}}$  form the stream of retrieval transactions that leave the SBS/RS. The inter-departure time distribution of this stream is denoted by  $D_{out,1}$  and its arrival rate is equal to  $\lambda_R$ .

Depending on the number of picking stations, the retrieval transactions that have a picking station as destination are routed to the first picking station with a probability of  $\frac{1}{n_{ps}}$ . Therefore, out of all retrieval transactions that leave the SBS/RS, a single transaction is routed to the first picking station with a probability of  $\frac{p_{ps}}{n_{ps}}$ . The inter-arrival time distribution of this stream is denoted by  $A_{ps,1}$ , whereas the inter-arrival time distribution of the stream bypassing the first picking station is denoted by  $A_{\bar{ps},1}$ . We assume that the service times of the picking stations are independent and identically distributed according to the random variable  $B_{ps}$ . The resulting inter-departure time distribution of the first picking station is denoted by  $D_{ps,1}$ . This stream includes both the stream of empty bins that leaves the system for replenishment and the stream of non-empty bins that directly re-enters the SBS/RS. Since all bins stay on the conveyor system, the departure rate is equal to the arrival rate at the picking station. Additionally, a stream of single items that are picked out of the bins is generated. This stream is of no interest within the scope of this thesis and leaves the system directly. Finally, the departure stream of the first picking station and the bypassing stream of the first picking station form a stream of transactions with an inter-departure time distribution  $D_{out,2}$ .

The procedure described for the first picking station repeats itself until the transaction stream after the last picking station is formed. The inter-departure time distribution of this stream is denoted by  $D_{out,n_{ps}+1}$  (if there are no picking stations,  $D_{out,1}$  leaves the system). The stream includes both storage transactions that re-enter the SBS/RS and retrieval transactions that leave the system. Hereby, the retrieval transactions include both the bins that are directly transported from the SBS/RS to the subsequent processes and the bins that are emptied at the picking stations. Therefore, they are routed out of the system with a probability of  $(1 - p_{ps} + p_{ps} \cdot p_{em})$ . The inter-departure time distribution of

this stream leaving the system is denoted by  $D_{out}$ . The other part of the stream is re-entering the SBS/RS with an inter-arrival time distribution of  $A_{S_{re-en}}$ . Since the probability to re-enter the SBS/RS is equal to  $(p_{ps} - p_{ps} \cdot p_{em})$ , the arrival rate of this stream is determined as follows.

$$\lambda_{S_{re-en}} = (p_{ps} - p_{ps} \cdot p_{em}) \cdot \lambda_R \quad (5.5)$$

The notation used in this section is presented in table 5.2. Before we can determine the network performance measures, we first have to model the service time distributions of the vehicle and lift stations. This will be shown in the next section.

Table 5.2: Notation of the random variables used to describe the distributions of the queueing model

---

$A_R$	Inter-arrival time distribution of the retrieval transactions
$A_S$	Inter-arrival time distribution of the storage transactions
$A_{S_{re-en}}$	Inter-arrival time distribution of the storage transactions that re-enter the system
$A_{S_{rep}}$	Inter-arrival time distribution of the storage transactions for replenishment
$A_{lift_{in}}$	Inter-arrival time distribution of the storage transactions at a lift <sub>in</sub> station
$A_{lift_{out}}$	Inter-arrival time distribution of the retrieval transactions at a lift <sub>out</sub> station
$A_{ps,j}$	Inter-arrival time distribution of the retrieval transactions at the $j$ -th picking station
$A_{\bar{ps},j}$	Inter-arrival time distribution of the transactions that are bypassing the $j$ -th picking station
$A_{veh}$	Inter-arrival time distribution of the storage and retrieval transactions at a vehicle station
$A_{veh,R}$	Inter-arrival time distribution of the retrieval transactions at a vehicle station
$A_{veh,S}$	Inter-arrival time distribution of the storage transactions at a vehicle station

---

$D_{lift_{in}}$	Inter-departure time distribution of the storage transactions at a lift <sub>in</sub> station
$D_{lift_{out}}$	Inter-departure time distribution of the retrieval transactions at a lift <sub>out</sub> station
$D_{out}$	Inter-departure time distribution of the retrieval transactions that are physically leaving the system
$D_{out,j}$	Inter-departure time distribution of the transactions before the split at the $j$ -th picking station
$D_{ps,j}$	Inter-departure time distribution of the transactions at the $j$ -th picking station
$D_{veh}$	Inter-departure time distribution of the storage and retrieval transactions at a vehicle station
$D_{veh,R}$	Inter-departure time distribution of the retrieval transactions at a vehicle station
$B_{lift_{in}}$	Service time distribution of a lift <sub>in</sub> station
$B_{lift_{out}}$	Service time distribution of a lift <sub>out</sub> station
$B_{ps}$	Service time distribution of a picking station
$B_{veh}$	Service time distribution of a vehicle station

---

### 5.1.3 Modeling of the service time distributions

#### Service time distribution of the vehicle station

Since the service time distribution of the vehicle depends on the physical dimensions of the tier, in which the vehicle is located, we first define the possible positions of the vehicle within the tier. We define the position of the vehicle within the tier by the horizontal position at column  $x$  and the vertical position at level  $yl$  of its load handling device. This results in the tuple  $(x, yl)$ , describing the position of the vehicle within the tier. As depicted in figure 5.3, the (un)loading points of a tier are located at  $x = -1$  and  $yl = 0$  ( $x = 0$  is designated for the first storage location of a tier). Moreover, the storage locations of a tier are located at  $(x, yl)$  with  $x \in \{0, 1, \dots, n_c - 1\}$  and  $yl \in \{0, 1, \dots, n_l - 1\}$ . In the second step, we describe the sequence of events, during which a transaction uses the vehicle. We denote this sequence of events as transaction cycle.

Afterward, we classify the possible transaction cycles. The cycle of a storage/retrieval transaction at the vehicle station consists of the following parts.

1. Vehicle travels from its actual position to the position where it loads the bin.
  - Storage transaction: loading point of the tier.
  - Retrieval transaction: random retrieval location.
2. Vehicle loads the bin.
3. Vehicle travels to the position where it unloads the bin.
  - Storage transaction: random storage location.
  - Retrieval transaction: unloading point of the tier.
4. Vehicle unloads the bin.

Therefore, we classify a transaction cycle by the type of transaction  $TT$ , which can be either a storage transaction  $S$  or a retrieval transaction  $R$ , the position of the vehicle when being requested  $H_v = (h_x, h_{yl})$  and the storage/retrieval location of the transaction  $\bar{H}_v = (\bar{h}_x, \bar{h}_{yl})$ . These three elements form the triple  $[TT, H_v, \bar{H}_v]_v$ , where

$$TT \in \{S, R\}, \quad (5.6)$$

$$\begin{aligned} H_v \in & \{(-1, 0), (h_x, h_{yl}) \mid h_x \in \{0, \dots, n_c - 1\} \\ & \wedge h_{yl} \in \{0, \dots, n_l - 1\}\}, \end{aligned} \quad (5.7)$$

$$\bar{H}_v \in \{(\bar{h}_x, \bar{h}_{yl}) \mid \bar{h}_x \in \{0, \dots, n_c - 1\} \wedge \bar{h}_{yl} \in \{0, \dots, n_l - 1\}\}. \quad (5.8)$$

Furthermore, we define the probability of a transaction cycle by  $P[TT, H_v, \bar{H}_v]_v$  and the respective service time by  $t[TT, H_v, \bar{H}_v]_v$ . Note that due to the discrete-time approach the service time may only assume multiples of the constant time increment  $t_{inc}$ .

Hence, we can determine the probability  $\beta_{veh,i}$  that the service time distribution  $B_{veh}$  assumes a value of  $i$  time increments. This is achieved by summing up the probabilities  $P[TT, H_v, \bar{H}_v]_v$  of those transaction cycles  $[TT, H_v, \bar{H}_v]_v$  that lead to a service time  $t[TT, H_v, \bar{H}_v]_v$  of  $i$  time increments.

$$\beta_{veh,i} = \sum_{\forall [TT, H_v, \bar{H}_v]_v \mid t[TT, H_v, \bar{H}_v]_v = i} P[TT, H_v, \bar{H}_v]_v \quad \forall i = 0 \dots i_{max} \quad (5.9)$$

If we are only interested in the distribution of service times of the retrieval transactions  $B_{veh,R}$ , we can determine the probability  $\beta_{veh,R,i}$  that  $B_{veh,R}$  assumes a value of  $i$  time increments by only considering the retrieval transactions. Thus, we sum up the probabilities  $P[R, H_v, \bar{H}_v]_v$  of those transaction cycles  $[R, H_v, \bar{H}_v]_v$  that lead to a service time  $t[R, H_v, \bar{H}_v]_v$  of  $i$  time increments. Since we do not consider the storage transactions, we have to normalize the distribution by dividing the probabilities  $P[R, H_v, \bar{H}_v]_v$  by the probability that the transaction is a retrieval transaction  $p_R$ .

$$\beta_{veh,R,i} = \sum_{\forall [R, H_v, \bar{H}_v]_v \mid t[R, H_v, \bar{H}_v]_v = i} \frac{P[R, H_v, \bar{H}_v]_v}{p_R} \quad \forall i = 0 \dots i_{max} \quad (5.10)$$

For the computation of the cycle probabilities  $P[TT, H_v, \bar{H}_v]_v$  and the respective cycle times  $t[TT, H_v, \bar{H}_v]_v$ , four cases can be distinguished.

1. Case  $[S, (-1, 0), \bar{H}_v]_v$ : in the first case, a storage transaction that uses the vehicle to store a bin at position  $\bar{H}_v$  requests a vehicle that is located at the (un)loading point.

The cycle probability is determined by three components. First, the probability that the transaction is of type storage, which is equal to  $(1 - p_R)$ . Second, the probability that the vehicle is located at the (un)loading point of the tier when being requested. Due to the POSC dwell point strategy, the vehicle idles at the (un)loading point after each retrieval transaction. Therefore, the probability that the vehicle is located at the (un)loading point when being requested is equal to  $p_R$ . Third, the probability to store the bin at position  $\bar{H}_v$ . Due to the random storage assignment rule, this probability is equal to  $\frac{1}{n_c n_l}$ . The product of these three components results in the cycle probability.

The cycle time is determined as follows. It consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel time  $t_v^{(-1,0),(\bar{h}_x,\bar{h}_{yl})}$  of the vehicle from the loading point at position  $(-1, 0)$  to the storage location at position  $(\bar{h}_x, \bar{h}_{yl})$ . In general, the travel times between any given positions  $(x, yl)$  and  $(x', yl')$  are denoted by  $t_v^{(x,yl),(x',yl')}$ . Due to the discrete-

time approach, we round the cycle time to the next time increment by dividing the sum of these times by the constant time increment  $t_{inc}$ , adding 0.5 and rounding down.

$$P[S, (-1, 0), \bar{H}_v]_v = (1 - p_R) \cdot p_R \cdot \frac{1}{n_c \cdot n_l} \quad \forall \bar{H}_v \quad (5.11)$$

$$t[S, (-1, 0), \bar{H}_v]_v = \left\lfloor \frac{2 \cdot t_{v,trans} + t_v^{(-1,0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 0.5 \right\rfloor \quad \forall \bar{H}_v \quad (5.12)$$

2. Case  $[S, H_v, \bar{H}_v]_v$ : in the second case, a storage transaction that uses the vehicle to store a bin at position  $\bar{H}_v$  requests a vehicle that is located at a position  $H_v$  inside the aisle, i.e. not at the (un)loading point.

The cycle probability is equal to the product of the probability that the transaction is of type storage, the probability that the vehicle is located at position  $H_v$  inside the aisle when being requested, and the probability to store the bin at position  $\bar{H}_v$ . As in the first case, the probability that the transaction is of type storage is equal to  $(1 - p_R)$  and the probability to store the bin at position  $\bar{H}_v$  is equal to  $\frac{1}{n_c \cdot n_l}$ . Since the vehicle only idles at a position inside the aisle when the last transaction was a storage transaction, the probability that the vehicle is located at position  $H_v$  inside the aisle when being requested is equal to the product of the following probabilities: the probability of a storage transaction  $(1 - p_R)$  and the probability that the vehicle idles at position  $H_v$  after completing the storage transaction, which is equal to  $\frac{1}{n_c \cdot n_l}$ .

The cycle time consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel times of the vehicle from position  $(h_x, h_{yl})$  to the loading point at position  $(-1, 0)$  and from  $(-1, 0)$  to the storage location at position  $(\bar{h}_x, \bar{h}_{yl})$ .

$$P[S, H_v, \bar{H}_v]_v = (1 - p_R)^2 \cdot \left( \frac{1}{n_c \cdot n_l} \right)^2 \quad (5.13)$$

$\forall H_v \setminus (-1, 0), \forall \bar{H}_v$

$$t[S, H_v, \bar{H}_v]_v = \left[ \frac{2 \cdot t_{v,trans} + t_v^{(h_x, h_{yl}), (-1, 0)}}{t_{inc}} \right. \\ \left. + \frac{t_v^{(-1, 0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 0.5 \right] \quad (5.14)$$

$\forall H_v \setminus (-1, 0), \forall \bar{H}_v$

3. Case  $[R, (-1, 0), \bar{H}_v]_v$ : in the third case, a retrieval transaction that uses the vehicle to retrieve a bin at position  $\bar{H}_v$  requests a vehicle that is located at the (un)loading point.

The cycle probability is equal to the product of the probability that the transaction is of type retrieval  $p_R$ , the probability that the vehicle is located at the (un)loading point of the tier when being requested, which is equal to  $p_R$ , and the probability to retrieve the bin at position  $\bar{H}_v$ , which is equal to  $\frac{1}{n_c \cdot n_l}$ .

The cycle time consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel times of the vehicle from the (un)loading point at position  $(-1, 0)$  to position  $(\bar{h}_x, \bar{h}_{yl})$  and the identical way back.

$$P[R, (-1, 0), \bar{H}_v]_v = p_R^2 \cdot \frac{1}{n_c \cdot n_l} \quad \forall \bar{H}_v \quad (5.15)$$

$$t[R, (-1, 0), \bar{H}_v]_v = \left[ \frac{2 \cdot t_{v,trans} + 2 \cdot t_v^{(-1, 0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 0.5 \right] \quad (5.16)$$

$\forall \bar{H}_v$

4. Case  $[R, H_v, \bar{H}_v]_v$ : in the fourth case, a retrieval transaction that uses the vehicle to retrieve a bin at position  $\bar{H}_v$  requests a vehicle that is located at position  $H_v$  inside the aisle, i.e. not at the (un)loading point.

The cycle probability is equal to the product of the probability that the transaction is of type retrieval  $p_R$ , the probability that the vehicle is located at position  $H_v$  inside the aisle when being requested, which is equal to  $(1 - p_R) \cdot \frac{1}{n_c \cdot n_l}$ , and the probability to retrieve the bin at position  $\bar{H}_v$ , which is equal to  $\frac{1}{n_c \cdot n_l}$ .

The cycle time consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel times of the vehicle from position  $(h_x, h_{yl})$  to position  $(\bar{h}_x, \bar{h}_{yl})$  and from position  $(\bar{h}_x, \bar{h}_{yl})$  to the unloading point at position  $(-1, 0)$ .

$$P[R, H_v, \bar{H}_v]_v = p_R \cdot (1 - p_R) \cdot \left( \frac{1}{n_c \cdot n_l} \right)^2 \quad (5.17)$$

$$\forall H_v \setminus (-1, 0), \forall \bar{H}_v$$

$$t[R, H_v, \bar{H}_v]_v = \left[ \frac{2 \cdot t_{v,trans} + t_v^{(h_x, h_{yl}), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + \frac{t_v^{(\bar{h}_x, \bar{h}_{yl}), (-1, 0)}}{t_{inc}} + 0.5 \right] \quad (5.18)$$

$$\forall H_v \setminus (-1, 0), \forall \bar{H}_v$$

The determination of the cycle times incorporates the travel times between any given positions  $(x, yl)$  and  $(x', yl')$ . Due to the simultaneous movement of the load handling device in horizontal and vertical directions, the travel time is equal to the maximum of the horizontal movement time  $t_{v,x}^{(x, x')}$  from position  $x$  to  $x'$  and the vertical movement time  $t_{v,y}^{(yl, yl')}$  from position  $yl$  to  $yl'$ .

$$t_v^{(x, yl), (x', yl')} = \max \{ t_{v,x}^{(x, x')}; t_{v,y}^{(yl, yl')} \} \quad (5.19)$$

The distance  $d_{v,x}^{(x, x')}$  between position  $x$  and  $x'$  is computed as follows.

$$d_{v,x}^{(x, x')} = |x - x'| \cdot d_c \quad (5.20)$$

Given the distance  $d_{v,x}^{(x,x')}$ , the vehicle's constant acceleration/deceleration rates  $a_{v,x}$ , and the maximum velocity  $v_{v,x}$ ,  $t_{v,x}^{(x,x')}$  can be determined depending on whether the needed distance to accelerate/decelerate the vehicle to/from the maximum velocity is shorter or greater than  $d_{v,x}^{(x,x')}$ . For a detailed description on how to compute the movement times of automated storage and retrieval systems, the reader is referred to Arnold and Furmans (2009).

$$t_{v,x}^{(x,x')} = \begin{cases} 2 \cdot \sqrt{\frac{d_{v,x}^{(x,x')}}{a_{v,x}}} & , d_{v,x}^{(x,x')} \leq \frac{v_{v,x}^2}{a_{v,x}} \\ \frac{d_{v,x}^{(x,x')}}{v_{v,x}} + \frac{v_{v,x}}{a_{v,x}} & , d_{v,x}^{(x,x')} > \frac{v_{v,x}^2}{a_{v,x}} \end{cases} \quad (5.21)$$

We use the same approach to determine the vertical distance  $d_{v,y}^{(yl,yl')}$  and the vertical movement time  $t_{v,y}^{(yl,yl')}$ .

$$d_{v,y}^{(yl,yl')} = |yl - yl'| \cdot d_l \quad (5.22)$$

$$t_{v,y}^{(yl,yl')} = \begin{cases} 2 \cdot \sqrt{\frac{d_{v,y}^{(yl,yl')}}{a_{v,y}}} & , d_{v,y}^{(yl,yl')} \leq \frac{v_{v,y}^2}{a_{v,y}} \\ \frac{d_{v,y}^{(yl,yl')}}{v_{v,y}} + \frac{v_{v,y}}{a_{v,y}} & , d_{v,y}^{(yl,yl')} > \frac{v_{v,y}^2}{a_{v,y}} \end{cases} \quad (5.23)$$

### Service time distribution of the lift<sub>in</sub> station

In analogy to the vehicle station, we first define the possible vertical positions of the incoming lift within the aisle by the parameter  $yt$ . As depicted in figure 5.3, the tiers are located at  $yt \in \{0, 1, \dots, n_t - 1\}$ . Furthermore, we denote the vertical position of the input point of the incoming lift by  $in$ .

In the following, we describe the sequence of events during which a storage transaction uses the incoming lift to transport the bin from the input point to the target tier.

1. Lift travels from its actual position to the input point of the lift.
2. Lift loads the bin.
3. Lift travels to the target tier.
4. Lift unloads the bin.

As before, we classify the possible transaction cycles by the type of transaction, which only can be a storage transaction  $S$ , the position of the lift when being requested  $H_{l_{in}} = h_{yt}$ , which can only be the position of one of the tiers due to the POSC dwell point strategy, and the position of the target tier  $\bar{H}_{l_{in}} = \bar{h}_{yt}$ . These three elements form the triple  $[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}}$ , where

$$H_{l_{in}} \in \{h_{yt} \mid h_{yt} \in \{0, \dots, n_t - 1\}\}, \quad (5.24)$$

$$\bar{H}_{l_{in}} \in \{\bar{h}_{yt} \mid \bar{h}_{yt} \in \{0, \dots, n_t - 1\}\}. \quad (5.25)$$

The probability of a transaction cycle is defined by  $P[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}}$  and the respective service time by  $t[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}}$ . In analogy to the vehicle station, we can determine the probability that the service time distribution  $B_{lift_{in}}$  assumes a value of  $i$  time increments.

$$\beta_{lift_{in}, i} = \sum_{\forall [S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}} \mid t[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}} = i} P[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}} \quad \forall i = 0 \dots i_{max} \quad (5.26)$$

Since all transactions using the incoming lift are storage transactions, only two components determine the cycle probabilities. The probability that the lift is located at position  $H_{l_{in}}$  when being requested and the probability to transport the bin to position  $\bar{H}_{l_{in}}$ . Due to the random storage assignment rule, both probabilities are equal to  $\frac{1}{n_t}$ .

The cycle time consists of the times to load and unload the bin  $t_{l_{in}, trans}$  as well as the travel times  $t_{l_{in}}^{(h_{yt}, in)}$  of the lift from position  $h_{yt}$  to the input point and  $t_{l_{in}}^{(in, \bar{h}_{yt})}$  from the input point to position  $\bar{h}_{yt}$ . In general,  $t_{l_{in}}^{(yt, yt')}$  denotes the travel time of the lift from position  $yt$  to position  $yt'$ .

$$P[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}} = \left( \frac{1}{n_t} \right)^2 \quad \forall H_{l_{in}}, \forall \bar{H}_{l_{in}} \quad (5.27)$$

$$t[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}} = \left\lfloor \frac{2 \cdot t_{l_{in}, trans} + t_{l_{in}}^{(h_{yt}, in)} + t_{l_{in}}^{(in, \bar{h}_{yt})}}{t_{inc}} + 0.5 \right\rfloor \quad (5.28)$$

$$\forall H_{l_{in}}, \forall \bar{H}_{l_{in}}$$

The distance  $d_{l_{in}}^{(yt, in)}$  between position  $yt$  and the input point is computed as follows.

$$d_{l_{in}}^{(yt, in)} = |yt \cdot d_t - d_{l_{in}}| \quad (5.29)$$

Given the distance  $d_{l_{in}}^{(yt, in)}$ , the incoming lift's constant acceleration/deceleration rates  $a_{l_{in}}$ , and the maximum velocity  $v_{l_{in}}$ ,  $t_{l_{in}}^{(yt, in)}$  can be determined depending on whether the needed distance to accelerate/decelerate the lift to/from the maximum velocity is shorter or greater than  $d_{l_{in}}^{(yt, in)}$ .

$$t_{l_{in}}^{(yt, in)} = \begin{cases} 2 \cdot \sqrt{\frac{d_{l_{in}}^{(yt, in)}}{a_{l_{in}}}} & , d_{l_{in}}^{(yt, in)} \leq \frac{v_{l_{in}}^2}{a_{l_{in}}} \\ \frac{d_{l_{in}}^{(yt, in)}}{v_{l_{in}}} + \frac{v_{l_{in}}}{a_{l_{in}}} & , d_{l_{in}}^{(yt, in)} > \frac{v_{l_{in}}^2}{a_{l_{in}}} \end{cases} \quad (5.30)$$

Obviously, we use the same approach to determine the distance  $d_{l_{in}}^{(in, yt)}$  and the time  $t_{l_{in}}^{(in, yt)}$ .

### Service time distribution of the lift<sub>out</sub> station

Similar to the incoming lift station, we define the possible vertical positions of the outgoing lift within the aisle by the parameter  $yt$ . As depicted in figure 5.3, the tiers are located at  $yt \in \{0, 1, \dots, n_t - 1\}$ . Furthermore, we denote the position of the output point of the outgoing lift by  $out$ .

The sequence of events, during which a retrieval transaction uses the outgoing lift to transport the bin from the tier it was retrieved to the output point, is as follows.

1. Lift travels from its actual position to the retrieval tier.
2. Lift loads the bin.
3. Lift travels to the output point.
4. Lift unloads the bin.

We classify the possible transaction cycles by the following three elements. First, by the type of transaction, which only can be a retrieval transaction  $R$  since no storage transaction is requesting the outgoing lift. Second, by the position of the lift when being requested  $H_{l_{out}} = out$ , which only can be at the output point of the aisle due to the POSC dwell point strategy. Third, by the position of the tier  $\bar{H}_{l_{out}} = \bar{h}_{yt}$  from where the bin is being retrieved. These three elements form the triple  $[R, out, \bar{H}_{l_{out}}]_{l_{out}}$ , where

$$\bar{H}_{l_{out}} \in \{\bar{h}_{yt} \mid \bar{h}_{yt} = \{0, \dots, n_t - 1\}\}. \quad (5.31)$$

The probability of a transaction cycle is defined by  $P[R, out, \bar{H}_{l_{out}}]_{l_{out}}$  and the respective service time by  $t[R, out, \bar{H}_{l_{out}}]_{l_{out}}$ . Hence, we can determine the probability that the service time distribution  $B_{lift_{out}}$  assumes a value of  $i$  time increments as follows.

$$\beta_{lift_{out}, i} = \sum_{\forall [R, out, \bar{H}_{l_{out}}]_{l_{out}} \mid t[R, out, \bar{H}_{l_{out}}]_{l_{out}} = i} P[R, out, \bar{H}_{l_{out}}]_{l_{out}} \quad \forall i = 0 \dots i_{max} \quad (5.32)$$

Since all transactions that use the outgoing lift are retrieval transactions and the point of service completion is always the output point, only the probability to retrieve the bin from position  $\bar{H}_{l_{out}}$  determines the cycle probabilities. Due to the equal access frequency among the tiers, the probability is equal to  $\frac{1}{n_t}$ .

The cycle time consists of the times to load and unload the bin  $t_{l_{out}, trans}$  as well as the travel times  $t_{l_{out}}^{(out, \bar{h}_{yt})}$  of the lift from the output point to position  $\bar{h}_{yt}$  and the identical way back.

$$P[R, out, \bar{H}_{l_{out}}]_{l_{out}} = \frac{1}{n_t} \quad \forall \bar{H}_{l_{out}} \quad (5.33)$$

$$t[R, out, \bar{H}_{l_{out}}]_{l_{out}} = \left\lfloor \frac{2 \cdot t_{l_{out}, trans} + 2 \cdot t_{l_{out}}^{(out, \bar{H}_{yt})}}{t_{inc}} + 0.5 \right\rfloor \quad \forall \bar{H}_{l_{out}} \quad (5.34)$$

The distance  $d_{l_{out}}^{(out, yt)}$  between the output point and a position  $yt$  is computed as follows.

$$d_{l_{out}}^{(out, yt)} = |yt \cdot d_t - d_{l_{out}}| \quad (5.35)$$

As before, given the outgoing lift's constant acceleration/deceleration rates  $a_{l_{out}}$  and the maximum velocity  $v_{l_{out}}$ ,  $t_{l_{out}}^{(out, yt)}$  can be determined depending on whether the needed distance to accelerate/decelerate the lift to/from the maximum velocity is shorter or greater than  $d_{l_{out}}^{(out, yt)}$ .

$$t_{l_{out}}^{(out, yt)} = \begin{cases} 2 \cdot \sqrt{\frac{d_{l_{out}}^{(out, yt)}}{a_{l_{out}}}} & , d_{l_{out}}^{(out, yt)} \leq \frac{v_{l_{out}}^2}{a_{l_{out}}} \\ \frac{d_{l_{out}}^{(out, yt)}}{v_{l_{out}}} + \frac{v_{l_{out}}}{a_{l_{out}}} & , d_{l_{out}}^{(out, yt)} > \frac{v_{l_{out}}^2}{a_{l_{out}}} \end{cases} \quad (5.36)$$

### 5.1.4 Computation of the network performance measures

The performance measures of interest are the utilization of the vehicles  $\rho_{veh}$ , the incoming lifts  $\rho_{lift_{in}}$ , the outgoing lifts  $\rho_{lift_{out}}$  and the pickers  $\rho_{ps}$ , as well as the distributions of the number of waiting bins at the arrival instant at the buffers in front of the incoming lifts  $Q_{lift_{in}}$  and the picking stations  $Q_{ps,1}, \dots, Q_{ps,n_{ps}}$ . Furthermore, we want to determine the distribution of the retrieval transaction time  $T_R$ , which is defined as the time that a retrieval transactions spends in the SBS/RS, i.e. the time span from entering the network until leaving the outgoing lift of the SBS/RS. It includes the waiting time for the vehicle to be available, the service time of the vehicle, the waiting time for the outgoing lift to be available, and the service time of the outgoing lift.

Given the transaction flow and the service time distributions, we can directly determine the utilization of the resources mentioned above. Since the stream of

incoming storage transactions is split equally among the incoming lift stations, the utilization of these stations is determined as follows.

$$\rho_{lift_{in}} = \frac{\lambda_S}{n_a} \cdot E[B_{lift_{in}}] \quad (5.37)$$

The same reasoning can be applied to the vehicle stations. Since the incoming streams of both the storage and the retrieval transactions are split equally among the vehicle stations, the transactions arrive at the vehicle station with an arrival rate of  $\frac{\lambda_S + \lambda_R}{n_a \cdot n_t}$ . Hence, the utilization is computed as follows.

$$\rho_{veh} = \frac{\lambda_S + \lambda_R}{n_a \cdot n_t} \cdot E[B_{veh}] \quad (5.38)$$

Regarding the outgoing lift station, the equal split of the retrieval transactions among the aisles leads to an arrival rate of the retrieval transactions at the outgoing lift station that is equal to  $\frac{\lambda_R}{n_a}$ . Thus, we determine the utilization of the outgoing lift station as follows.

$$\rho_{lift_{out}} = \frac{\lambda_R}{n_a} \cdot E[B_{lift_{out}}] \quad (5.39)$$

Due to the equal split of the retrieval transactions that are sent to one of the picking stations among those stations, the arrival rate of the retrieval transactions at the picking stations is equal to  $\frac{p_{ps} \cdot \lambda_R}{n_{ps}}$ . Therefore, we calculate the utilization of a picking station, which is identical for all picking stations, as follows.

$$\rho_{ps} = \frac{p_{ps} \cdot \lambda_R}{n_{ps}} \cdot E[B_{ps}] \quad (5.40)$$

In order to determine the network performance measures  $T_R$ ,  $Q_{lift_{in}}$ , and  $Q_{ps,1}, \dots, Q_{ps,n_{ps}}$ , we decompose the network into independent  $G|G|1$  queueing systems. As stated before in section 4.2.1, given the inter-arrival time distribution  $A_j$  and the service time distribution  $B_j$  of a queueing system  $j$ , we can determine both its waiting time distribution  $W_j$  and its inter-departure time distribution  $D_j$  using the algorithms of Grassmann and Jain (1989) and Jain and Grassmann (1988), respectively.

Because of the random storage assignment rule, the equal access frequency of the retrieval transactions, and the identical service time distributions, the inter-arrival time distributions of all vehicle stations will be identical. This in turn leads to identical waiting time distributions of the customers at all vehicle stations. The same applies to the incoming and outgoing lift stations. Therefore, the sojourn time distribution of a retrieval transaction at a vehicle station can be computed as follows. Since both storage and retrieval transactions are served at the vehicle stations, we determine the sojourn time distribution of the retrieval transactions  $T_{veh,R}$  by the convolution of the waiting time distribution  $W_{veh}$  and the retrieval service time distribution  $B_{veh,R}$ .

$$T_{veh,R} = W_{veh} \otimes B_{veh,R} \quad (5.41)$$

Since only retrieval transactions are served by the outgoing lift station, we determine the sojourn time distribution of the retrieval transactions  $T_{lift_{out}}$  by the convolution of the waiting time distribution  $W_{lift_{out}}$  and the service time distribution  $B_{lift_{out}}$ .

$$T_{lift_{out}} = W_{lift_{out}} \otimes B_{lift_{out}} \quad (5.42)$$

Finally, we determine the retrieval transaction time distribution  $T_R$  by the convolution of the sojourn time distribution of the retrieval transactions at the vehicle station  $T_{veh,R}$  and the sojourn time distribution at the outgoing lift station  $T_{lift_{out}}$ .

$$T_R = T_{veh,R} \otimes T_{lift_{out}} \quad (5.43)$$

Moreover, we can determine for a given queueing system  $j$  the number of customers  $N_j$  at the arrival instant of a customer using the algorithm of Furmans and Zillus (1996). Given  $N_j$ , where  $\eta_{j,i}$  denotes the probability that there are  $i$  customers at the queueing system  $j$  at the arrival instant, we can directly derive the number of waiting customers  $Q_j$  at the arrival instant, where  $q_{j,i}$  denotes the probability that there are  $i$  customers waiting at the arrival instant.

$$q_{j,i} = \begin{cases} \eta_{j,0} + \eta_{j,1} & i = 0 \\ \eta_{j,i+1} & i \neq 0 \end{cases} \quad \forall i = 0, \dots, i_{max} \quad (5.44)$$

The notation used in this section to describe the distributions of the performance measures is presented in table 5.3.

Table 5.3: Notation of the random variables used to describe the distributions of the performance measures

$B_{veh,R}$	Retrieval service time distribution of a vehicle station
$Q_{lift_{in}}$	Queue length distribution of a lift <sub>in</sub> station at the arrival instant
$Q_{ps,j}$	Queue length distribution of the $j$ -th picking station at the arrival instant
$T_{lift_{out}}$	Sojourn time distribution of the retrieval transactions at a lift <sub>out</sub> station
$T_R$	Retrieval transaction time distribution
$T_{veh,R}$	Sojourn time distribution of the retrieval transactions at a vehicle station
$W_{lift_{out}}$	Waiting time distribution of the retrieval transactions at a lift <sub>out</sub> station
$W_{veh}$	Waiting time distribution of the storage and retrieval transactions at a vehicle station

In a network with a linear topology, we can use the inter-departure time distribution of the upstream queueing system as the inter-arrival time distribution of the downstream queueing system. In contrast, to determine the inter-arrival time distribution of a downstream queueing system in a network with a non-linear topology, we have to split and/or merge customer streams of the upstream queueing systems. In our case, we use the methods described in section 4.2 to perform the split and merge operations.

Due to the re-entering stream of storage transactions (see figure 5.5), we determine the steady-state performance measures of the network using an iterative algorithm. In this algorithm, we use the computed distribution of the re-entering storage transaction stream of the  $(n_{it} - 1)$ -th iteration as input for the computation of the distribution of the storage transaction stream entering the SBS/RS in the  $n_{it}$ -th iteration. Since the inter-arrival time and inter-departure time distributions of all vehicle stations will be identical during the  $n_{it}$ -th iteration, we only need to consider one vehicle station during the iterations. The

same applies to the incoming and outgoing lift stations. Given that, the  $n_{it}$ -th iteration of the algorithm consists of the following steps.

1. Compute the inter-arrival time distribution  $A_{liftin}^{(n_{it})}$  of the incoming lift stations by
  - first merging  $A_{S_{re-en}}^{(n_{it}-1)}$  and  $A_{S_{rep}}$  to compute  $A_S^{(n_{it})}$ ,
  - and then splitting  $A_S^{(n_{it})}$  into  $n_a$  identical distributions  $A_{liftin}^{(n_{it})}$ .
2. Compute the inter-departure time distribution  $D_{liftin}^{(n_{it})}$  of the incoming lift stations.
3. Compute the inter-arrival time distribution  $A_{veh}^{(n_{it})}$  of the vehicle stations by
  - first splitting  $D_{liftin}^{(n_{it})}$  into  $n_t$  identical distributions  $A_{veh,S}^{(n_{it})}$ ,
  - and then merging  $A_{veh,S}^{(n_{it})}$  and  $A_{veh,R}$  to compute  $A_{veh}^{(n_{it})}$ .
4. Compute the sojourn time distribution  $T_{veh,R}^{(n_{it})}$  of the retrieval transactions at the vehicle station by
  - first computing the waiting time distribution  $W_{veh}^{(n_{it})}$  of the vehicle station,
  - and then computing  $T_{veh,R}^{(n_{it})}$  by the convolution of  $W_{veh}^{(n_{it})}$  and  $B_{veh,R}$ .
5. Compute the inter-departure time distribution  $D_{veh}^{(n_{it})}$  of the vehicle stations.
6. Compute the inter-arrival time distribution  $A_{liftout}^{(n_{it})}$  of the outgoing lift stations by
  - first splitting  $D_{veh}^{(n_{it})}$  into  $D_{veh,R}^{(n_{it})}$  and the stream of fulfilled storage transactions leaving the system,
  - and then merging  $n_t$  distributions  $D_{veh,R}^{(n_{it})}$  to compute  $A_{liftout}^{(n_{it})}$ .
7. Compute the sojourn time distribution  $T_{liftout}^{(n_{it})}$  of the retrieval transactions at the outgoing lift station by
  - first computing the waiting time distribution  $W_{liftout}^{(n_{it})}$  of the outgoing lift station,
  - and then computing  $T_{liftout}^{(n_{it})}$  by the convolution of  $W_{liftout}^{(n_{it})}$  and  $B_{liftout}$ .

8. Compute the inter-departure time distribution  $D_{lift_{out}}^{(n_{it})}$  of the outgoing lift stations.
9. Compute the inter-departure time distribution  $D_{out,1}^{(n_{it})}$  leaving the SBS/RS by merging  $n_a$  distributions  $D_{lift_{out}}^{(n_{it})}$ .
10. For all  $j = 1 \dots n_{ps}$  picking stations:
  - Compute the inter-arrival time distribution  $A_{ps,j}^{(n_{it})}$  of the  $j$ -th picking station by splitting  $D_{out,j}^{(n_{it})}$  into  $A_{ps,j}^{(n_{it})}$  and  $A_{\bar{ps},j}^{(n_{it})}$ .
  - Compute the inter-departure time distribution  $D_{ps,j}^{(n_{it})}$  of the  $j$ -th picking station.
  - Compute the inter-departure time distribution  $D_{out,j+1}^{(n_{it})}$  after the merge of  $A_{\bar{ps},j}^{(n_{it})}$  and  $D_{ps,j}^{(n_{it})}$ .
11. Split the inter-departure time distribution  $D_{out,n_{ps}+1}^{(n_{it})}$  to compute  $D_{out}^{(n_{it})}$  and  $A_{S_{re-en}}^{(n_{it})}$ .
12. Compute the retrieval transaction time distribution  $T_R^{(n_{it})}$  by the convolution of  $T_{veh,R}^{(n_{it})}$  and  $T_{lift_{out}}^{(n_{it})}$ .
13. If the absolute difference between the expected values of  $T_R^{(n_{it})}$  and  $T_R^{(n_{it}-1)}$  is bigger than a given  $\varepsilon_{it}$ , continue with iteration  $(n_{it} + 1)$  and use  $A_{S_{re-en}}^{(n_{it})}$  as input for the computation of the inter-arrival time distribution of the incoming lift stations in the  $(n_{it} + 1)$ -th iteration (step 1). Otherwise, stop the iterations and compute the steady-state distributions of the performance measures by
  - setting  $T_R = T_R^{(n_{it})}$ ,
  - setting  $D_{out} = D_{out}^{(n_{it})}$ ,
  - computing  $Q_{lift_{in}}$  given  $A_{lift_{in}}^{(n_{it})}$  and  $B_{lift_{in}}$ ,
  - computing  $Q_{ps,j}$  given  $A_{ps,j}^{(n_{it})}$  and  $B_{ps}$  for all  $j = 1 \dots n_{ps}$  picking stations.

We initialize the algorithm by splitting the retrieval inter-arrival time distribution  $A_R$  into  $n_a \cdot n_t$  identical distributions  $A_{veh,R}$ . Since these distributions do

not change during the iterations, we will use them as input for all iterations  $n_{it}$ . Furthermore, we start with an initial value of the retrieval transaction time distribution of  $E[T_R^{(n_{it}=0)}] = 0$ .

In the first iteration  $n_{it} = 1$ , we do not consider the storage transaction stream. Therefore, we start the first iteration with the computation of the performance measures of the vehicle station (step 4), and use  $A_{veh,R}$  as the inter-arrival time distribution, i.e.  $A_{veh}^{(n_{it}=1)} = A_{veh,R}$ . Beginning with the second iteration, we start the iterations with the computation of the inter-arrival time distribution  $A_{liftin}^{(n_{it})}$  of the incoming lift stations (step 1).

The detailed pseudo-code is presented in algorithm 1. The use of one of the methods to generate  $W_j$ ,  $D_j$ , or  $Q_j$  of queueing system  $j$  is denoted by  $G|G|1(A_j, B_j)$ . Moreover, the use of the split operation to determine the distribution  $A_\xi$  after the split of the distribution  $A_j$  in two or more directions is denoted by  $\text{Split}(A_j, p_\xi)$ , where  $p_\xi$  is the probability that a customer is routed to direction  $\xi$ . The use of the merge operation to merge two different distributions  $A_\xi$  and  $A_\zeta$  is denoted by  $\text{Merge}(A_\xi, A_\zeta)$ . Furthermore, the use of the merge operation to merge  $n_{Merge}$  identical distributions  $A_\xi$  is denoted by  $\text{Merge}(A_\xi, n_{Merge}\text{-times})$ .

---

**Algorithm 1** Determination of the performance measures of the tier-captive SBS/RS

```
1: procedure TIER-CAPTIVE( $n_a$ ,  $n_t$ ,  $n_{ps}$ ,  $A_R$ ,  $A_{S_{rep}}$ ,  $B_{liftin}$ ,  $B_{veh}$ ,  $B_{veh,R}$ ,  
     $B_{liftout}$ ,  $B_{ps}$ ,  $p_R$ ,  $p_{em}$ ,  $p_{ps}$ ,  $\varepsilon_{it}$ )  
2:    $n_{it} \leftarrow 0$   
3:    $E[T_R^{(n_{it})}] \leftarrow 0$   
4:   Generate  $A_{veh,R} \leftarrow \text{Split}(A_R, \frac{1}{n_a n_t})$   
5:   repeat  
6:      $n_{it} \leftarrow n_{it} + 1$   
7:     if  $n_{it} = 1$  then  
8:        $A_{veh}^{(n_{it})} \leftarrow A_{veh,R}$   
9:     else if  $n_{it} > 1$  then  
10:      Generate  $A_S^{(n_{it})} \leftarrow \text{Merge}(A_{S_{re-en}}^{(n_{it}-1)}, A_{S_{rep}})$   
11:      Generate  $A_{liftin}^{(n_{it})} \leftarrow \text{Split}(A_S^{(n_{it})}, \frac{1}{n_a})$   
12:      Generate  $D_{liftin}^{(n_{it})} \leftarrow G|G|1(A_{liftin}^{(n_{it})}, B_{liftin})$ 
```

---

```

13:           Generate  $A_{veh,S}^{(n_{it})} \leftarrow \text{Split}(D_{lift_{in}}^{(n_{it})}, \frac{1}{n_t})$ 
14:           Generate  $A_{veh}^{(n_{it})} \leftarrow \text{Merge}(A_{veh,S}^{(n_{it})}, A_{veh,R})$ 
15:       end if
16:       Generate  $W_{veh}^{(n_{it})} \leftarrow G|G|1(A_{veh}^{(n_{it})}, B_{veh})$ 
17:        $T_{veh,R}^{(n_{it})} \leftarrow W_{veh}^{(n_{it})} \otimes B_{veh,R}$ 
18:       Generate  $D_{veh}^{(n_{it})} \leftarrow G|G|1(A_{veh}^{(n_{it})}, B_{veh})$ 
19:       Generate  $D_{veh,R}^{(n_{it})} \leftarrow \text{Split}(D_{veh}^{(n_{it})}, p_R)$ 
20:       Generate  $A_{liftout}^{(n_{it})} \leftarrow \text{Merge}(D_{veh,R}^{(n_{it})}, n_t\text{-times})$ 
21:       Generate  $W_{liftout}^{(n_{it})} \leftarrow G|G|1(A_{liftout}^{(n_{it})}, B_{liftout})$ 
22:        $T_{liftout}^{(n_{it})} \leftarrow W_{liftout}^{(n_{it})} \otimes B_{liftout}$ 
23:       Generate  $D_{liftout}^{(n_{it})} \leftarrow G|G|1(A_{liftout}^{(n_{it})}, B_{liftout})$ 
24:        $j \leftarrow 1$ 
25:       Generate  $D_{out,j}^{(n_{it})} \leftarrow \text{Merge}(D_{liftout}^{(n_{it})}, n_a\text{-times})$ 
26:       if  $n_{ps} > 0$  then
27:           repeat
28:               Generate  $A_{ps,j}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, \frac{p_{ps}}{n_{ps}})$ 
29:               Generate  $A_{ps,j}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, 1 - \frac{p_{ps}}{n_{ps}})$ 
30:               Generate  $D_{ps,j}^{(n_{it})} \leftarrow G|G|1(A_{ps,j}^{(n_{it})}, B_{ps})$ 
31:               Generate  $D_{out,j+1}^{(n_{it})} \leftarrow \text{Merge}(A_{ps,j}^{(n_{it})}, D_{ps,j}^{(n_{it})})$ 
32:                $j \leftarrow j + 1$ 
33:           until  $j > n_{ps}$ 
34:       end if
35:       Generate  $D_{out}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, 1 - p_{ps} + p_{ps} \cdot p_{em})$ 
36:       Generate  $A_{S_{re-en}}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, p_{ps} - p_{ps} \cdot p_{em})$ 
37:        $T_R^{(n_{it})} \leftarrow T_{veh,R}^{(n_{it})} \otimes T_{liftout}^{(n_{it})}$ 
38:       until  $|E[T_R^{(n_{it})}] - E[T_R^{(n_{it}-1)}]| < \varepsilon_{it}$ 
39:        $T_R \leftarrow T_R^{(n_{it})}$ 
40:        $D_{out} \leftarrow D_{out}^{(n_{it})}$ 
41:       Generate  $Q_{liftin} \leftarrow G|G|1(A_{liftin}^{(n_{it})}, B_{liftin})$ 
42:       if  $n_{ps} > 0$  then

```

```
43:       $j \leftarrow 1$ 
44:      repeat
45:          Generate  $Q_{ps,j} \leftarrow G|G|1(A_{ps,j}^{(n_{it})}, B_{ps})$ 
46:           $j \leftarrow j + 1$ 
47:      until  $j > n_{ps}$ 
48:      end if
49:      return  $T_R, D_{out}, Q_{liftin}, Q_{ps,1}, \dots, Q_{ps,n_{ps}}$ 
50: end procedure
```

---

## 5.2 Tier-to-tier configuration

In analogy to section 5.1, we present the calculation of the performance measures of a tier-to-tier configuration. At first, we give a description of the system under consideration. Afterward, we determine the queueing model of the investigated system. The service time distributions are modeled in section 5.2.3. Finally, we present the algorithm to compute the performance measures.

### 5.2.1 System description

The system under investigation is a tier-to-tier SBS/RS. An example of such a system with three aisles and two picking stations is depicted in figure 5.6. In analogy to the tier-captive SBS/RS presented in section 5.1, it consists of  $n_a$  aisles,  $n_t$  tiers per aisle,  $n_l$  levels per tier and  $n_c$  storage columns on either side of the aisle. Again, the distances between two tiers, two levels of a tier and two storage columns are given by  $d_t$ ,  $d_l$  and  $d_c$ , respectively. As shown in figures 5.7 and 5.8, the storage columns of any level are equally sized and can hold one bin. Since the vehicles can not change the aisle, the system does not have any cross-aisles.

In contrast to the tier-captive configuration in section 5.1, in each aisle there is exactly one shuttle that can use the lift in front of the aisle to change the tiers of the aisle. Therefore, a shuttle is able to access all storage locations of an aisle. As in section 5.1, the shuttles have one load handling device with a capacity of one bin, and are able to serve all  $n_l$  levels of a tier. They travel simultaneously along the aisle and lift/lower the load handling device in vertical direction to-

wards the target level. Their maximum velocities in horizontal direction along the aisle and vertical direction to lift/lower the load handling device are  $v_{v,x}$  and  $v_{v,y}$ , respectively. The acceleration rates, which are identical to the deceleration rates in the respective direction, are given by  $a_{v,x}$  and  $a_{v,y}$ . The vehicle transfer times  $t_{v,trans}$  to load/unload the bin between the input/output points of the aisle and the shuttle are identical to the transfer times to load/unload the bin between the storage rack and the shuttle.

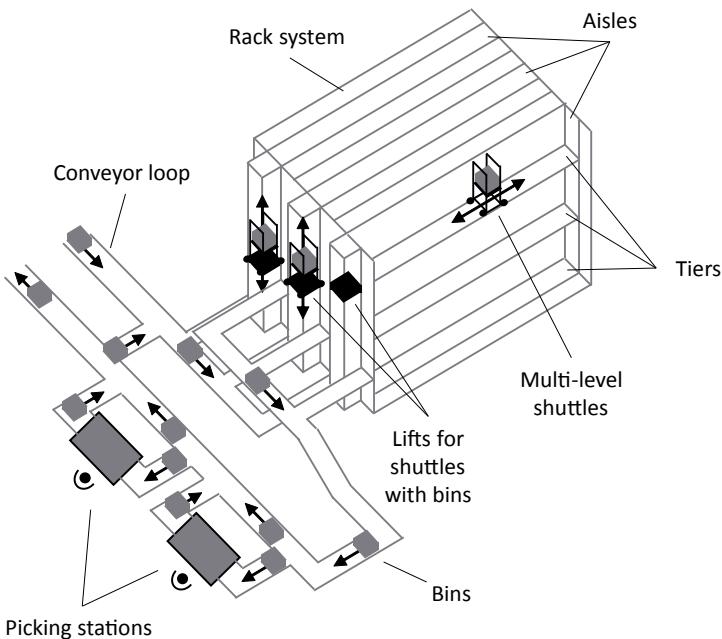


Figure 5.6: Tier-to-tier system with 3 aisles and 2 picking stations

Unlike the tier-captive configuration in section 5.1, directly in front of the aisle, one lift provides the vertical transport of the shuttles (with/without bin). The shuttles can access a tier of an aisle by driving out of the lift into the aisle. Therefore, the (un)loading points of a tier are located in front of the aisle at the position of the lift. In vertical direction, the distance between the (un)loading

points and the lowest level of a tier is 0. In horizontal direction, the distance between the (un)loading points and the first column of a tier is equal to  $d_c$ . In order to access the lift from within the aisle, the shuttle drives out of the aisle onto the lift. From there, the lift transports the shuttle to its target position, i.e. another tier or the input/output points of the aisle. At the input and output points of the aisle, the bins can be transferred from the incoming conveyor system to the shuttle or from the shuttle to the outgoing conveyor system. Each lift can hold one shuttle (with/without bin) and is only able to travel in vertical direction. Their maximum velocities in vertical direction are  $v_l$ . The acceleration rates, which are identical to the deceleration rates of the respective lift, are given by  $a_l$ . As in section 5.1, the distances between the lowest tier of the system and the input and output points of the aisles are given by  $d_{l_{in}}$  and  $d_{l_{out}}$ , respectively.

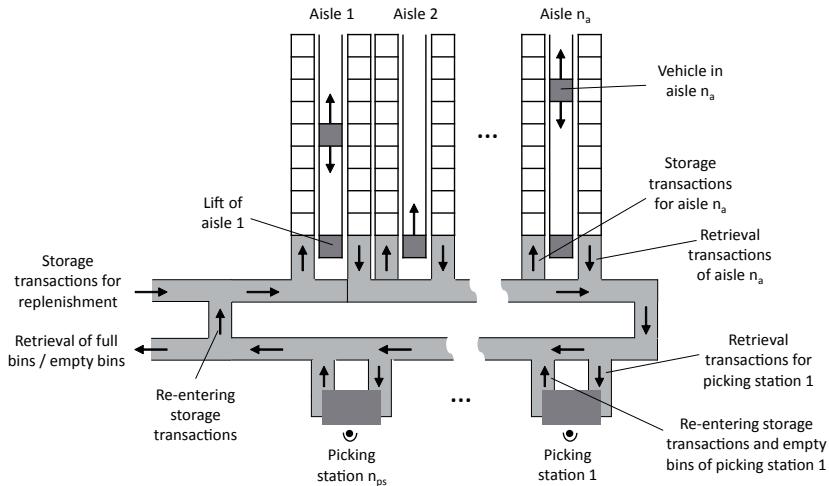


Figure 5.7: Plan view of the tier-to-tier system

Outside of the SBS/RS, a conveyor system transports the outgoing bins in the same way as described in section 5.1 from the output points of the aisles to their target destination such as one of the  $n_{ps}$  picking stations or another subsequent process. Therefore, the probability that a bin is routed to one of the picking

stations is denoted by  $p_{ps}$ . The same conveyor system transports the empty bins from the picking station to another process in the warehouse that refills the bins. The probability that a bin is emptied at the picking station is denoted by  $p_{em}$ . The conveyor system also transports the re-entering bins from the picking stations and the replenishment bins from the refilling process to the lifts. At each picking station, there is one worker that picks the required number of items out of the bin.

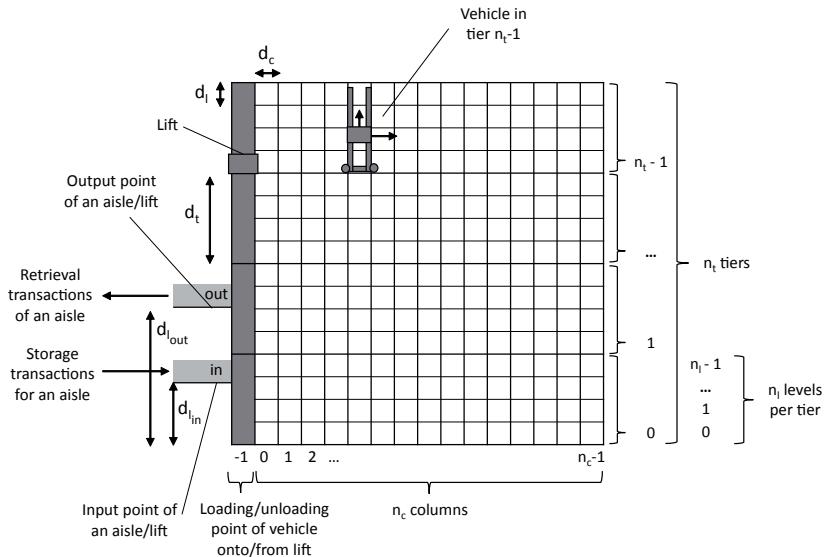


Figure 5.8: Aisle view of the tier-to-tier system

The control strategies of the system are equivalent to the control strategies of the tier-captive system in section 5.1. Due to the tier-to-tier configuration with only one shuttle per aisle and lift, the shuttles, which seize the lifts during the fulfillment of the transaction, operate in a single command cycle mode. They process the arriving storage and retrieval transactions based on a first-come-first-serve strategy. Furthermore, the incoming storage transactions are assigned randomly among the storage locations. The access frequency of the

retrieval transactions is shared equally among the storage locations. Also the probability to route a bin to one of the picking stations is shared equally among the picking stations. As a result, the system load is – on average – shared equally between the shuttles, lifts and picking stations. In general, both the shuttles and the lifts idle at the point of service completion.

Identical to the tier-captive system described in section 5.1, there are two types of transactions, namely retrieval and storage transactions. The retrieval transactions can both represent a transport order for a bin that needs to be transported to a picking station and a bin that needs to be directly transported to another subsequent process. Additionally, there are retrieval transactions for bins that got empty at the picking station and need to leave the system to be refilled. In contrast to that, the storage transactions both represent the transport orders for the re-entering bins after the picking process and the replenishment bins coming from the refilling process.

In analogy to figure 5.4, the flow of the transactions through the tier-to-tier system is represented by a graph model (see figure 5.9). The nodes represent the primary resources inside the SBS/RS as well as the picking stations outside the SBS/RS. The edges represent the flow of transactions between the primary resources inside the SBS/RS. As explained before, edges represent the conveyor system outside the SBS/RS. The lifts, the shuttles and the pickers are the secondary resources that lead to a population constraint at the primary resources. In figure 5.9, dotted rectangular shapes around the primary resources depict the parts of the system that have a population constraint of one due to the service by a single secondary resource (e.g., by a vehicle, lift or picker).

The dashed edges of the graph model represent the flow of retrieval transactions, whereas the solid edges represent the flow of storage transactions. A retrieval transaction uses both the shuttle and the lift to transport the requested bin out of the rack system to the outgoing conveyor system. In contrary to the tier-captive system in section 5.1, the shuttle is now able to use the lift to change the tiers. Thus, if the shuttle idles at the output point of the aisle on the lift, the retrieval transaction uses the lift to transport the shuttle to the retrieval tier. Afterward, it first uses the shuttle to retrieve the bin and then uses the lift to transport the shuttle with the bin to the output point, where the shuttle transfers the bin to the conveyor system.

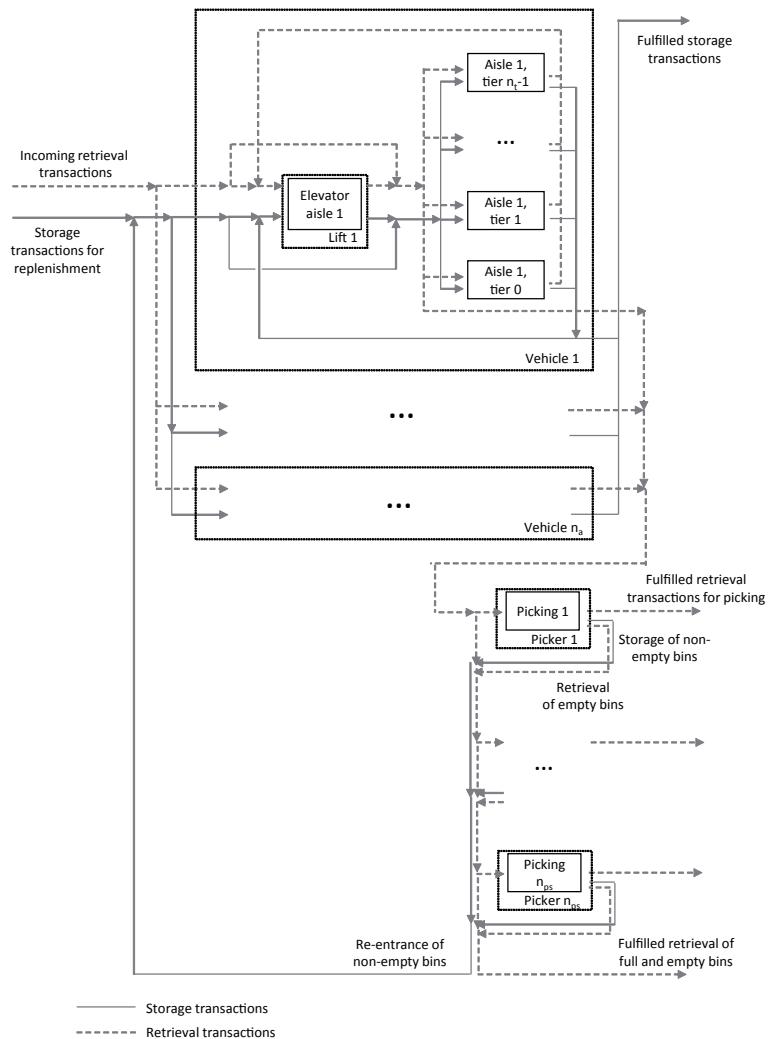


Figure 5.9: Graph model of the tier-to-tier system

If the shuttle is located within the aisle when being requested, the retrieval transaction uses the shuttle, which itself uses the lift to change tiers, to retrieve the bin. If the bin is located in the same tier as the shuttle when being requested, the retrieval transaction orders the shuttle to travel to the retrieval location, retrieve the bin and travel to the lift. At the lift, it orders the lift to travel to the output point, where the shuttle releases the bin. If the bin is located at another tier, the retrieval transaction orders the shuttle to travel to the lift and the lift to transfer the shuttle to the retrieval tier. In the retrieval tier, it orders the shuttle to retrieve the bin and to travel to the lift. Finally, it orders the lift to travel to the output point, where the shuttle releases the bin.

Outside the SBS/RS, the flow of storage and retrieval transactions uses the connected conveyor system and the picking stations in the same way as in tier-captive systems. If the bin needs to be directly transported to a subsequent process, the retrieval transaction uses the conveyor system to leave the system, bypassing all picking stations. If the bin needs to be transported to a picking station, the system routes the bin using the shortest path on the conveyor system to the picking station that requested the bin. At the picking station, the retrieval transaction is continued by a picker who takes the requested number of items out of the bin. At that point, the retrieval transaction is fulfilled and leaves the system. After the picking process, the system generates a new retrieval transaction if the bin is empty. It uses the conveyor system to leave the system, bypassing all picking stations. If the bin is not empty, the system generates a new storage transaction with a target storage location that is defined by the random storage assignment rule. It uses the shortest path on the conveyor system to transport the bin to the input point of the target aisle. Also the storage transactions that enter the system for replenishment use the conveyor system to route the bins to the input point of the target aisle.

Inside the SBS/RS, both types of storage transactions use the shuttle and the lift to store the bin in the rack system. If the shuttle idles at the output point of the aisle on the lift, the storage transaction uses the lift to transport the shuttle first to the input point, where it loads the bin, and then to the storage tier, where it uses the shuttle to store the bin. If the shuttle is located within the aisle when being requested, the storage transaction uses the shuttle, which itself uses the lift to change tiers, to store the bin. The storage transaction first orders the shuttle to travel to the lift and the lift to transfer the shuttle to the input point, where it loads the bin. Afterward, it orders the lift to transport the shuttle to the

storage tier and the shuttle to store the bin at the target destination. After that, the storage transactions are fulfilled and virtually leave the system.

Table 5.4: Notation used in the system description

$a_l$	Lift acceleration/deceleration rate in vertical direction
$a_{v,x}$	Vehicle acceleration/deceleration rate in horizontal direction
$a_{v,y}$	Vehicle/LHD acceleration/deceleration rate in vertical direction
$n_a$	Number of aisles
$n_c$	Number of columns on either side of an aisle
$n_l$	Number of levels per tier
$n_t$	Number of tiers per aisle
$n_{ps}$	Number of picking stations
$p_{em}$	Probability that a bin is empty after picking
$p_{ps}$	Probability that the target of the retrieval transaction is a picking station
$t_{v,trans}$	Vehicle transfer time
$v_l$	Lift velocity in vertical direction
$v_{v,x}$	Vehicle velocity in horizontal direction
$v_{v,y}$	Vehicle/LHD velocity in vertical direction
$d_c$	Distance between two storage columns
$d_l$	Distance between two storage levels
$d_t$	Distance between two storage tiers
$d_{l_{in}}$	Distance between the lowest tier and the input point of the lift
$d_{l_{out}}$	Distance between the lowest tier and the output point of the lift

### 5.2.2 Queueing network model

The discrete-time queueing network that is used to model tier-to-tier SBS/RSSs is depicted in figure 5.10. In this network, the parts of the system with a population constraint of one are represented by single server queueing systems with unlimited queueing capacities. Since the parts of the system that are served by the vehicles represent parts of the system with a population constraint of one, we model these parts by queueing systems. In contrast to the queueing

network of the tier-captive system in section 5.1.2, we therefore represent the whole aisles by queueing systems. As in section 5.1.2, the pickers are represented by queueing systems as well.

As before, we assume that the transactions form one customer class and that the stochastic routing of the customers is determined by the routing matrix  $Z$ . The inter-arrival times of the incoming retrieval transactions are assumed to be independent and identically distributed according to the random variable  $A_R$ . Hence, the arrival rate of the retrieval transactions is determined as follows.

$$\lambda_R = \frac{1}{E[A_R]} \quad (5.45)$$

Due to the equal access frequency of the retrieval transactions to the storage locations, the routing probability of the retrieval transaction stream to one of the aisle stations, which incorporate all operations of the vehicle and the lift, is equal to  $\frac{1}{n_a}$ . The split of the retrieval transaction stream into  $n_a$  identical streams leads to the inter-arrival time distribution  $A_{aisle,R}$  of the retrieval transactions arriving at each aisle station.

The inter-arrival times of the storage transactions for replenishment are assumed to be independent and identically distributed according to the random variable  $A_{S_{rep}}$ . Thus, the arrival rate is computed as follows.

$$\lambda_{S_{rep}} = \frac{1}{E[A_{S_{rep}}]} \quad (5.46)$$

Together with the re-entering storage transactions, which have an inter-arrival time distribution of  $A_{S_{re-en}}$  and which have the arrival rate  $\lambda_{S_{re-en}} = \frac{1}{E[A_{S_{re-en}}]}$ , they form the stream of the incoming storage transactions. In analogy to section 5.1.2, the inter-arrival time distribution of this stream is denoted by  $A_S$ . Its arrival rate is determined as follows.

$$\lambda_S = \lambda_{S_{rep}} + \lambda_{S_{re-en}} \quad (5.47)$$

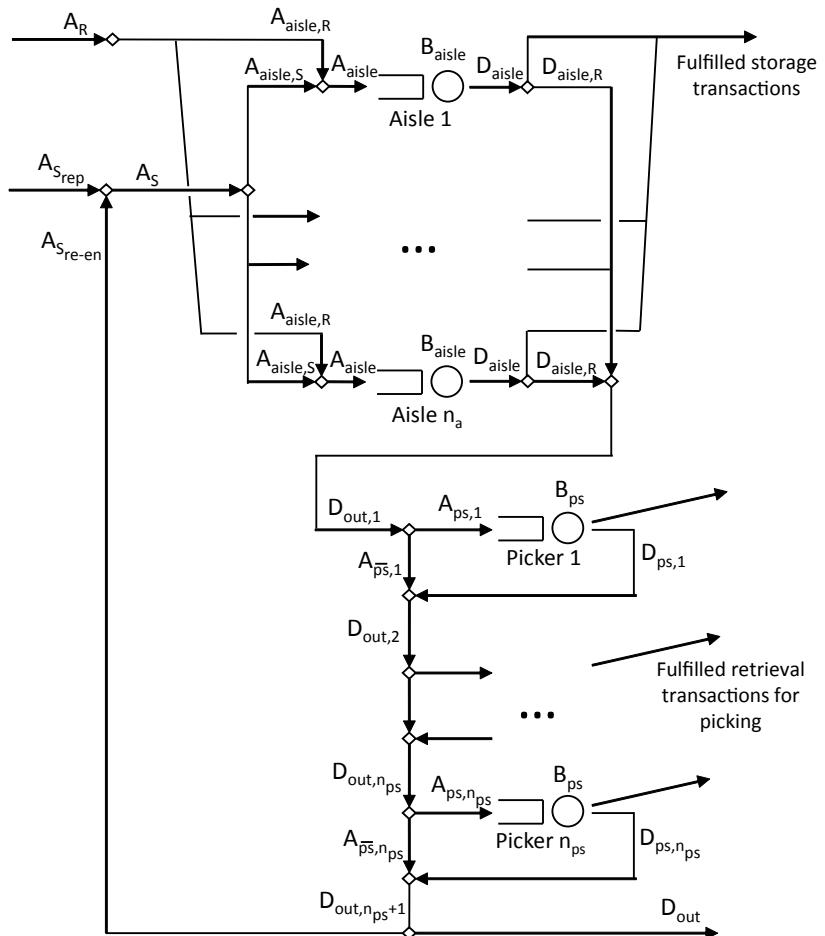


Figure 5.10: Queueing model of the tier-to-tier system

Due to the random storage assignment rule, the routing probability of the storage transaction stream to one of the aisle stations is equal to  $\frac{1}{n_a}$ . The split of the storage transaction stream into  $n_a$  identical streams leads to the inter-arrival time distribution  $A_{aisle,S}$  of the storage transactions arriving at each aisle station. Together with the retrieval transactions, they form the stream of arriving transactions at an aisle station. The inter-arrival time distribution of this stream, which is identical for each aisle station, is denoted by  $A_{aisle}$ . Due to the identical physical design and control policies for all lifts and vehicles, the service time distributions  $B_{aisle}$ , which we assume to be i.i.d., are identical for each aisle station. Therefore, also the inter-departure time distributions  $D_{aisle}$  are identical for each aisle station.

Only the retrieval transaction streams are routed to the picking stations or to other subsequent processes. Hence, after each aisle station the retrieval transactions are split from the total departure stream. Due to the equal distribution of incoming storage and retrieval transactions among the aisle stations, the ratio of retrieval transactions among all departing transactions of an aisle station is equal to  $p_R$ .

$$p_R = \frac{\lambda_R}{\lambda_S + \lambda_R} \quad (5.48)$$

This ratio is equal to the routing probability of the retrieval transaction stream of an aisle station to the picking stations or to other subsequent processes. The inter-departure time distribution of this stream is denoted by  $D_{aisle,R}$ . The other part of the departure stream of an aisle station represents the fulfilled storage transactions that virtually leave the system after storing the bins in the rack. All inter-departure time distributions form the stream of retrieval transactions that leave the SBS/RS. As in section 5.1.2, the inter-departure time distribution of this stream is denoted by  $D_{out,1}$  and its arrival rate is equal to  $\lambda_R$ .

Following that, the part of the queueing network describing the picking process and the routing of the customers on the conveyor system is identical to the equivalent part of the queueing network of the tier-captive system presented in section 5.1.2.

The notation used in this section is presented in table 5.5. In the following section, we show how to model the service time distribution of the aisle station.

Table 5.5: Notation of the random variables used to describe the distributions of the queueing model

$A_R$	Inter-arrival time distribution of the retrieval transactions
$A_S$	Inter-arrival time distribution of the storage transactions
$A_{S_{re-en}}$	Inter-arrival time distribution of the storage transactions that re-enter the system
$A_{S_{rep}}$	Inter-arrival time distribution of the storage transactions for replenishment
$A_{aisle}$	Inter-arrival time distribution of the storage and retrieval transactions at an aisle station
$A_{aisle,R}$	Inter-arrival time distribution of the retrieval transactions at an aisle station
$A_{aisle,S}$	Inter-arrival time distribution of the storage transactions at an aisle station
$A_{ps,j}$	Inter-arrival time distribution of the retrieval transactions at the $j$ -th picking station
$A_{\bar{ps},j}$	Inter-arrival time distribution of the transactions that are bypassing the $j$ -th picking station
$B_{aisle}$	Service time distribution of an aisle station
$B_{ps}$	Service time distribution of a picking station
$D_{aisle}$	Inter-departure time distribution of the storage and retrieval transactions at an aisle station
$D_{aisle,R}$	Inter-departure time distribution of the retrieval transactions at an aisle station
$D_{out}$	Inter-departure time distribution of the transactions that are leaving the system
$D_{out,j}$	Inter-departure time distribution of the transactions before the split into the $j$ -th picking station
$D_{ps,j}$	Inter-departure time distribution of the transactions at the $j$ -th picking station

### 5.2.3 Modeling of the service time distributions

The modeling of the service time distribution of the aisle station of the tier-to-tier system is performed in the same way as the modeling of the service time distributions of the vehicle,  $\text{lift}_{in}$ , and  $\text{lift}_{out}$  stations of the tier-captive system in section 5.1.3.

Since the service time distribution of the aisle depends on the physical dimensions of the aisle, in which the vehicle is located, we first define the possible positions of the vehicle within the aisle. In contrast to section 5.1.3, the vehicle is able to change the tiers of the aisle. Therefore, we define the position of the vehicle within the aisle by the horizontal position at column  $x$ , the vertical position at level  $yl$  of its load handling device, and additionally the vertical position at tier  $yt$ . This results in the triple  $(x, yt, yl)$ , describing the position of the vehicle within the aisle. As depicted in figure 5.8, the (un)loading points of a vehicle onto/from the lift at a tier are located at  $x = -1$ ,  $yl = 0$  and  $yt \in \{0, 1, \dots, n_t - 1\}$ . The input point of the lift is located at  $(-1, yt = in, 0)$ , whereas the output point is located at  $(-1, yt = out, 0)$ . Moreover, the storage locations are located at  $(x, yt, yl)$  with  $x \in \{0, 1, \dots, n_c - 1\}$ ,  $yt \in \{0, 1, \dots, n_t - 1\}$  and  $yl \in \{0, 1, \dots, n_l - 1\}$ .

In the next step, we describe the sequence of events during which a transaction uses the vehicle. Afterward, we classify the possible transaction cycles. The cycle of a storage/retrieval transaction at the aisle station consists of the following parts.

1. Vehicle travels (with/without using the lift) from its actual position to the position where it loads the bin.
  - Storage transaction: input point of the aisle.
  - Retrieval transaction: random retrieval location.
2. Vehicle loads the bin.
3. Vehicle travels (by using the lift) to the position where it unloads the bin.
  - Storage transaction: random storage location.
  - Retrieval transaction: output point of the aisle.
4. Vehicle unloads the bin.

Thus, a transaction cycle is classified by the type of transaction  $TT$ , which can be either a storage transaction  $S$  or a retrieval transaction  $R$ , the position of the vehicle when being requested  $H_a = (h_x, h_{yt}, h_{yl})$  and the storage/retrieval location of the transaction  $\bar{H}_a = (\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl})$ . In analogy to section 5.1.3, these three elements form the triple  $[TT, H_a, \bar{H}_a]_a$ , where

$$TT \in \{S, R\}, \quad (5.49)$$

$$\begin{aligned} H_a \in & \{(-1, out, 0), (h_x, h_{yt}, h_{yl}) \mid h_x \in \{0, \dots, n_c - 1\} \\ & \wedge h_{yt} \in \{0, \dots, n_t - 1\} \wedge h_{yl} \in \{0, \dots, n_l - 1\}\}, \end{aligned} \quad (5.50)$$

$$\begin{aligned} \bar{H}_a \in & \{(\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl}) \mid \bar{h}_x \in \{0, \dots, n_c - 1\} \\ & \wedge \bar{h}_{yt} \in \{0, \dots, n_t - 1\} \wedge \bar{h}_{yl} \in \{0, \dots, n_l - 1\}\}. \end{aligned} \quad (5.51)$$

To determine the service time distribution  $B_{aisle}$ , we define the probability of a transaction cycle by  $P[TT, H_a, \bar{H}_a]_a$  and the respective service time by  $t[TT, H_a, \bar{H}_a]_a$ . As in the sections before, the service time may only assume multiples of the constant time increment  $t_{inc}$ .

Hence, in analogy to section 5.1.3, we can determine the probability  $\beta_{aisle,i}$  that  $B_{aisle}$  assumes a value of  $i$  time increments by summing up the probabilities  $P[TT, H_a, \bar{H}_a]_a$  of those transaction cycles  $[TT, H_a, \bar{H}_a]_a$  that lead to a service time  $t[TT, H_a, \bar{H}_a]_a$  of  $i$  time increments.

$$\beta_{aisle,i} = \sum_{\forall [TT, H_a, \bar{H}_a]_a \mid t[TT, H_a, \bar{H}_a]_a = i} P[TT, H_a, \bar{H}_a]_a \quad \forall i = 0 \dots i_{max} \quad (5.52)$$

In order to compute the service time distribution of the retrieval process  $B_{aisle,R}$ , we consider only the retrieval transactions. Thus, we sum up the probabilities  $P[R, H_a, \bar{H}_a]_a$  of those transaction cycles  $[R, H_a, \bar{H}_a]_a$  that lead to a service time  $t[R, H_a, \bar{H}_a]_a$  of  $i$  time increments. Afterward, we normalize the distribution by dividing the probabilities  $P[R, H_a, \bar{H}_a]_a$  by the probability that the transaction is a retrieval transaction  $p_R$ .

$$\beta_{aisle,R,i} = \sum_{\forall [R, H_a, \bar{H}_a]_a \mid t[R, H_a, \bar{H}_a]_a = i} \frac{P[R, H_a, \bar{H}_a]_a}{p_R} \quad \forall i = 0 \dots i_{max} \quad (5.53)$$

For the computation of the cycle probabilities  $P[TT, H_a, \bar{H}_a]_a$  and the respective cycle times  $t[TT, H_a, \bar{H}_a]_a$ , we distinguish different cases and determine both the cycle probabilities and the cycle times in analogy to section 5.1.3. In the tier-to-tier system under investigation, four cases can be distinguished.

1. Case  $[S, (-1, out, 0), \bar{H}_a]_a$ : in the first case, a storage transaction that uses the vehicle and the lift to store a bin at position  $\bar{H}_a$  requests a vehicle that is located at the output point of the aisle.

The cycle probability is determined by three components. First, the probability that the transaction is of type storage, which is equal to  $(1 - p_R)$ . Second, the probability that the vehicle is located at the output point when being requested. Due to the POSC dwell point strategy, the vehicle idles at the output point after each retrieval transaction. Therefore, the probability that the vehicle is located at the output point when being requested is equal to  $p_R$ . Third, the probability to store the bin at position  $\bar{H}_a$ . Due to the random storage assignment rule, this probability is equal to  $\frac{1}{n_c \cdot n_t \cdot n_l}$ . The product of these three components results in the cycle probability.

The cycle time consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel times  $t_l^{(out,in)}$  of the lift from the output point to the input point,  $t_l^{(in,\bar{h}_{yt})}$  of the lift from the input point to the target tier and  $t_v^{(-1,0),(\bar{h}_x,\bar{h}_{yl})}$  of the vehicle from the (un)loading point of the tier to the storage location of the bin. As in the sections before, we round the cycle time to the next time increment by dividing the sum of these times by the constant time increment  $t_{inc}$ , adding 0.5 and rounding down.

$$P[S, (-1, out, 0), \bar{H}_a]_a = (1 - p_R) \cdot p_R \cdot \frac{1}{n_c \cdot n_t \cdot n_l} \quad \forall \bar{H}_a \quad (5.54)$$

$$t[S, (-1, out, 0), \bar{H}_a]_a = \left\lfloor \frac{2 \cdot t_{v,trans} + t_l^{(out,in)} + t_l^{(in,\bar{h}_{yt})}}{t_{inc}} + \frac{t_v^{(-1,0),(\bar{h}_x,\bar{h}_{yl})}}{t_{inc}} + 0.5 \right\rfloor \quad \forall \bar{H}_a \quad (5.55)$$

2. Case  $[S, H_a, \bar{H}_a]_a$ : in the second case, a storage transaction that uses the vehicle and the lift to store a bin at position  $\bar{H}_a$  requests a vehicle that is located at position  $H_a$ .

The cycle probability is equal to the product of the probability  $(1 - p_R)$  that the transaction is of type storage, the probability that the vehicle is located at position  $H_a$  when being requested, and the probability  $\frac{1}{n_c \cdot n_t \cdot n_l}$  to store the bin at position  $\bar{H}_a$ . Since the vehicle only idles at a position inside the aisle when the last transaction was a storage transaction, the probability that the vehicle is located at position  $H_a$  when being requested is equal to  $(1 - p_R) \cdot \frac{1}{n_c \cdot n_t \cdot n_l}$ .

The cycle time consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel times of the vehicle from position  $(h_x, h_{yt}, h_{yl})$  to the (un)loading point of the tier at position  $(-1, h_{yt}, 0)$ , the travel times of the lift from position  $h_{yt}$  to the input position  $in$  and from the input position  $in$  to position  $\bar{h}_{yt}$ , and the vehicle travel time from the (un)loading point of the tier  $(-1, \bar{h}_{yt}, 0)$  to the storage location at position  $(\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl})$ .

$$P[S, H_a, \bar{H}_a]_a = (1 - p_R)^2 \cdot \left( \frac{1}{n_c \cdot n_t \cdot n_l} \right)^2 \quad (5.56)$$

$$\forall H_a \setminus (-1, out, 0), \forall \bar{H}_a$$

$$t[S, H_a, \bar{H}_a]_a = \left[ \frac{2 \cdot t_{v,trans} + t_v^{(h_x, h_{yl}), (-1, 0)} + t_l^{(h_{yt}, in)}}{t_{inc}} \right. \\ \left. + \frac{t_l^{(in, \bar{h}_{yt})} + t_v^{(-1, 0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 0.5 \right] \quad (5.57)$$

$$\forall H_a \setminus (-1, out, 0), \forall \bar{H}_a$$

3. Case  $[R, (-1, out, 0), \bar{H}_a]_a$ : in the third case, a retrieval transaction that uses the vehicle and the lift to retrieve a bin at position  $\bar{H}_a$  requests a vehicle that is located at the output point of the aisle.

The cycle probability is equal to the product of the probability  $p_R$  that the transaction is of type retrieval, the probability  $p_R$  that the vehicle

is located at the output point when being requested, and the probability  $\frac{1}{n_c \cdot n_t \cdot n_l}$  to retrieve the bin at position  $\bar{H}_a$ .

The cycle time consists of the times to load and unload the bin  $t_{v,trans}$  as well as the travel times of the lift from the output point at position  $out$  to position  $\bar{h}_{yt}$  and the way back, and the travel times of the vehicle from the (un)loading point of the tier at position  $(-1, \bar{h}_{yt}, 0)$  to position  $(\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl})$  and the way back.

$$P[R, (-1, out, 0), \bar{H}_a]_a = p_R^2 \cdot \frac{1}{n_c \cdot n_t \cdot n_l} \quad \forall \bar{H}_a \quad (5.58)$$

$$t[R, (-1, out, 0), \bar{H}_a]_a = \left[ \begin{array}{l} \frac{2 \cdot t_{v,trans} + 2 \cdot t_l^{(out, \bar{h}_{yt})}}{t_{inc}} \\ + \frac{2 \cdot t_v^{(-1, 0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 0.5 \end{array} \right] \quad \forall \bar{H}_a \quad (5.59)$$

4. Case  $[R, H_a, \bar{H}_a]_v$ : in the fourth case, a retrieval transaction that uses the vehicle and the lift to retrieve a bin at position  $\bar{H}_a$  requests a vehicle that is located at position  $H_a$ .

The cycle probability is equal to the product of the probability  $p_R$  that the transaction is of type retrieval, the probability  $(1 - p_R) \cdot \frac{1}{n_c \cdot n_t \cdot n_l}$  that the vehicle is located at position  $H_a$  when being requested, and the probability  $\frac{1}{n_c \cdot n_t \cdot n_l}$  to retrieve the bin at position  $\bar{H}_a$ .

If the requested vehicle is located at the same tier as the bin that will be retrieved, i.e.  $h_{yt} = \bar{h}_{yt}$ , the cycle time consists of the following times: the times to load and unload the bin  $t_{v,trans}$ , the travel times of the vehicle from position  $(h_x, h_{yt}, h_{yl})$  to position  $(\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl})$  and from position  $(\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl})$  to the (un)loading point of the tier at position  $(-1, \bar{h}_{yt}, 0)$ , as well as the travel time of the lift from position  $\bar{h}_{yt}$  to the output point. If the requested vehicle is not located at the same tier as the bin that will be retrieved, the cycle time consists of the following times: the times to load and unload the bin  $t_{v,trans}$ , the travel time of the vehicle from position  $(h_x, h_{yt}, h_{yl})$  to the (un)loading point of the tier at position  $(-1, h_{yt}, 0)$ , the time of the lift to travel from position  $h_{yt}$  to  $\bar{h}_{yt}$ , the travel times of

the vehicle from the (un)loading point of the tier at position  $(-1, \bar{h}_{yt}, 0)$  to position  $(\bar{h}_x, \bar{h}_{yt}, \bar{h}_{yl})$  and the way back, as well as the travel time of the lift from position  $\bar{h}_{yt}$  to the output point.

$$P[R, H_a, \bar{H}_a]_a = p_R \cdot (1 - p_R) \cdot \left( \frac{1}{n_c \cdot n_t \cdot n_l} \right)^2 \quad (5.60)$$

$$t[R, H_a, \bar{H}_a]_a = \begin{cases} \left\lfloor \frac{\frac{2 \cdot t_{v,trans} + t_v^{(h_x, h_{yl}), (\bar{h}_x, \bar{h}_{yl}), (-1, 0)}}{t_{inc}} + t_v^{(\bar{h}_x, \bar{h}_{yl}), (\bar{h}_x, \bar{h}_{yl}), (-1, 0)}}{t_{inc}} + 0.5 \right\rfloor & , h_{yt} = \bar{h}_{yt} \\ \left\lfloor \frac{\frac{2 \cdot t_{v,trans} + t_v^{(h_x, h_{yl}), (-1, 0)} + t_l^{(h_{yt}, \bar{h}_{yt}), (-1, 0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 2 \cdot t_v^{(-1, 0), (\bar{h}_x, \bar{h}_{yl})}}{t_{inc}} + 0.5 \right\rfloor & , h_{yt} \neq \bar{h}_{yt} \end{cases} \quad (5.61)$$

$$\forall H_a \setminus (-1, out, 0), \forall \bar{H}_a$$

The travel times of the vehicle  $t_v^{(x, yt), (x', yt')}$  between any given positions  $(x, yt, yt')$  and  $(x', yt', yt')$  within the same tier, i.e.  $yt = yt'$ , as well as the travel times of the lift  $t_l^{(yt, yt')}$  between any given positions  $yt$  and  $yt'$  are determined in the same way as in section 5.1.3.

### 5.2.4 Computation of the network performance measures

In the tier-to-tier system, the performance measures of interest are

- the utilization of the aisles  $\rho_{aisle}$ ,
- the utilization of the vehicles  $\rho_{veh}$  (which is identical to  $\rho_{aisle}$ ),
- the utilization of the pickers  $\rho_{ps}$ ,
- the distribution of the number of waiting bins at the arrival instant at the buffers in front of the lifts  $Q_{aisle,S}$  as well as

- the distribution of the number of waiting bins at the arrival instant at the buffers in front of the picking stations  $Q_{ps,1}, \dots, Q_{ps,n_{ps}}$ .

If we would like to compute the utilization of the lift, we also need to compute the average time that the vehicle uses the lift during the transaction cycle. Furthermore, as in section 5.1.4, we want to determine

- the distribution of the retrieval transaction time  $T_R$ ,

which is defined as the total time that a retrieval transactions spends in the SBS/RS, i.e. the time span from entering the network until leaving the lift of the SBS/RS at the output point of the aisle. It includes the waiting time for the vehicle to be available and the service time of the aisle, which includes the time using the vehicle and the lift.

Given the transaction flow and the service time distributions, we can directly determine the utilization of the resources mentioned above. Since the incoming streams of both the storage and the retrieval transactions are split equally among the aisle stations, the transactions arrive at the aisle station with an arrival rate of  $\frac{\lambda_S + \lambda_R}{n_a}$ . Hence, the utilization is computed as follows.

$$\rho_{aisle} = \frac{\lambda_S + \lambda_R}{n_a} \cdot E[B_{aisle}] \quad (5.62)$$

As in section 5.1.4, the retrieval transactions that are sent to one of the picking stations are equally shared among those stations. Therefore, the arrival rate of the retrieval transactions at the picking stations is equal to  $\frac{p_{ps} \cdot \lambda_R}{n_{ps}}$ . We calculate the utilization of the picking station, which is identical for all picking stations, as follows.

$$\rho_{ps} = \frac{p_{ps} \cdot \lambda_R}{n_{ps}} \cdot E[B_{ps}] \quad (5.63)$$

We use the same approach as in section 5.1.4 to determine the network performance measures. Hence, we compute  $T_R$ ,  $Q_{aisle,S}$  and  $Q_{ps,1}, \dots, Q_{ps,n_{ps}}$  by decomposing the network into independent  $G|G|1$  queueing systems. As before, we use the algorithms of Grassmann and Jain (1989) and Jain and Grassmann (1988) to determine the waiting time distribution  $W_j$  and the inter-departure time distribution  $D_j$  of a given queueing system  $j$ .

Because of the random storage assignment rule, the equal access frequency of the retrieval transactions and the identical service time distributions, the inter-arrival time distributions and the waiting time distributions of all aisle stations will be identical. Therefore, the retrieval transaction time distribution  $T_R$  can be computed as follows. Since both storage and retrieval transactions are served at the aisle stations, we determine  $T_R$  by the convolution of the waiting time distribution  $W_{aisle}$  and the retrieval service time distribution  $B_{aisle,R}$ .

$$T_R = W_{aisle} \otimes B_{aisle,R} \quad (5.64)$$

Moreover, we can determine for a given queueing system  $j$  the number of customers  $N_j$  at the arrival instant of a customer using the algorithm of Furmans and Zillus (1996). As in the tier-captive system, we can directly derive the number of waiting customers  $Q_j$  at the arrival instant given  $N_j$  (see equation 5.44). Note that the queue consists of storage and retrieval transactions although only the storage transactions physically exist. In order to derive the number of storage transactions waiting  $Q_{j,S}$ , where  $q_{j,S,i}$  denotes the probability of  $i$  waiting storage transactions, we use the Binomial distribution. The notation used in this section to describe the distributions of the performance measures is presented in table 5.6.

$$q_{j,S,i} = \sum_{l=i}^{Q_{j,max}} q_{j,l} \cdot \binom{l}{i} \cdot \left( \frac{\lambda_S}{\lambda_S + \lambda_R} \right)^i \cdot \left( \frac{\lambda_R}{\lambda_S + \lambda_R} \right)^{l-i} \quad (5.65)$$

$$\forall i = 0 \dots Q_{j,S,max}$$

If we need to split and/or merge customer streams of the upstream queueing systems to determine the inter-arrival time distribution of a downstream queueing system, we use the methods described in section 4.2.

In a tier-to-tier configuration, we have a re-entering stream of storage transactions as well (see figure 5.10). Thus, we determine the steady-state performance measures in analogy to section 5.1.4 using an iterative algorithm. As before, we use the computed re-entering storage transaction stream of the  $(n_{it} - 1)$ -th iteration as input for the computation of the storage transaction stream entering the SBS/RS in the  $n_{it}$ -th iteration.

Table 5.6: Notation of the random variables used to describe the distributions of the performance measures

$B_{aisle,R}$	Retrieval service time distribution of an aisle station
$Q_{aisle}$	Queue length distribution of an aisle station at the arrival instant
$Q_{aisle,S}$	Storage transaction queue length distribution of an aisle station at the arrival instant
$Q_{ps,j}$	Queue length distribution of the $j$ -th picking station at the arrival instant
$T_R$	Retrieval transaction time distribution
$W_{aisle}$	Waiting time distribution of the storage and retrieval transactions at an aisle station

Since the inter-arrival time and inter-departure time distributions of all aisle stations will be identical during the  $n_{it}$ -th iteration, we only need to consider one aisle station during the iterations. Given that, the  $n_{it}$ -th iteration of the algorithm consists of the following steps.

1. Compute the inter-arrival time distribution  $A_{aisle}^{(n_{it})}$  of the aisle stations by
  - first merging  $A_{S_{re-en}}^{(n_{it}-1)}$  and  $A_{S_{rep}}$  to compute  $A_S^{(n_{it})}$ ,
  - then splitting  $A_S^{(n_{it})}$  into  $n_a$  identical distributions  $A_{aisle,S}^{(n_{it})}$ ,
  - and finally merging  $A_{aisle,S}^{(n_{it})}$  and  $A_{aisle,R}$  to compute  $A_{aisle}^{(n_{it})}$ .
2. Compute the waiting time distribution  $W_{aisle}^{(n_{it})}$  of the aisle stations.
3. Compute the inter-departure time distribution  $D_{aisle}^{(n_{it})}$  of the aisle stations.
4. Compute the inter-departure time distribution  $D_{out,1}^{(n_{it})}$  leaving the SBS/RS by
  - first splitting  $D_{aisle}^{(n_{it})}$  into  $D_{aisle,R}^{(n_{it})}$  and the stream of fulfilled storage transactions leaving the system,
  - and then merging  $n_a$  distributions  $D_{aisle,R}^{(n_{it})}$  to compute  $D_{out,1}^{(n_{it})}$ .
5. For all  $j = 1 \dots n_{ps}$  picking stations:

- Compute the inter-arrival time distribution  $A_{ps,j}^{(n_{it})}$  of the  $j$ -th picking station by splitting  $D_{out,j}^{(n_{it})}$  into  $A_{ps,j}^{(n_{it})}$  and  $A_{\bar{ps},j}^{(n_{it})}$ .
  - Compute the inter-departure time distribution  $D_{ps,j}^{(n_{it})}$  of the  $j$ -th picking station.
  - Compute the inter-departure time distribution  $D_{out,j+1}^{(n_{it})}$  after the merge of  $A_{\bar{ps},j}^{(n_{it})}$  and  $D_{ps,j}^{(n_{it})}$ .
6. Split the inter-departure time distribution  $D_{out,n_{ps}+1}^{(n_{it})}$  to compute  $D_{out}^{(n_{it})}$  and  $A_{S_{re-en}}^{(n_{it})}$ .
  7. Compute the retrieval transaction time distribution  $T_R^{(n_{it})}$  by the convolution of  $W_{aisle}^{(n_{it})}$  and  $B_{aisle,R}$ .
  8. If the absolute difference between the expected values of  $T_R^{(n_{it})}$  and  $T_R^{(n_{it}-1)}$  is bigger than a given  $\varepsilon_{it}$ , continue with iteration  $(n_{it} + 1)$  and use  $A_{S_{re-en}}^{(n_{it})}$  as input for the computation of the inter-arrival time distribution of the aisle stations in the  $(n_{it} + 1)$ -th iteration (step 1). Otherwise, stop the iterations and compute the steady-state distributions of the performance measures by
    - setting  $T_R = T_R^{(n_{it})}$ ,
    - setting  $D_{out} = D_{out}^{(n_{it})}$ ,
    - computing  $Q_{aisle,S}$  given  $A_{aisle}^{(n_{it})}$ ,  $B_{aisle}$  and  $Q_{aisle}$ ,
    - computing  $Q_{ps,j}$  given  $A_{ps,j}^{(n_{it})}$  and  $B_{ps}$  for all  $j = 1 \dots n_{ps}$  picking stations.

We initialize the algorithm by splitting the retrieval inter-arrival time distribution  $A_R$  into  $n_a$  identical distributions  $A_{aisle,R}$ . Since these distributions do not change during the iterations, we will use them as input for all iterations  $n_{it}$ . Furthermore, we start with an initial value of the retrieval transaction time distribution of  $E[T_R^{(n_{it}=0)}] = 0$ .

In the first iteration  $n_{it} = 1$ , we do not consider the storage transaction stream. Therefore, we start the first iteration using  $A_{aisle,R}$  as the inter-arrival time distribution of the aisle station, i.e.  $A_{aisle}^{(n_{it}=1)} = A_{aisle,R}$ . Beginning with the second

iteration, we start the iterations with the computation of the inter-arrival time distribution  $A_{aisle}^{(n_{it})}$  of the aisle stations.

The detailed pseudo-code is given by algorithm 2. As before, the use of one of the methods to generate  $W_j$ ,  $D_j$ , or  $Q_j$  of queueing system  $j$  is denoted by  $G|G|1(A_j, B_j)$ . The derivation of the distribution of waiting storage transactions  $Q_{j,S}$  in a queue with both storage and retrieval transactions according to equation 5.65 is denoted by  $\text{Binomial}(Q_j, 1 - p_R)$ . Moreover, as in section 5.1.4, the use of the split operation to determine the distribution  $A_\xi$  after the split of the distribution  $A_j$  in two or more directions is denoted by  $\text{Split}(A_j, p_\xi)$ , where  $p_\xi$  is the probability that a customer is routed to direction  $\xi$ . The use of the merge operation to merge two different distributions  $A_\xi$  and  $A_\zeta$  is denoted by  $\text{Merge}(A_\xi, A_\zeta)$ . Furthermore, the use of the merge operation to merge  $n_{Merge}$  identical distributions  $A_\xi$  is denoted by  $\text{Merge}(A_\xi, n_{Merge}\text{-times})$ .

---

**Algorithm 2** Determination of the performance measures of the tier-to-tier SBS/RS

---

```
1: procedure TIER-TO-TIER( $n_a, n_{ps}, A_R, A_{S_{rep}}, B_{aisle}, B_{aisle,R}, B_{ps}, p_R, p_{em},$ 
   $p_{ps}, \varepsilon_{it}$ )
2:    $n_{it} \leftarrow 0$ 
3:    $E[T_R^{(n_{it})}] \leftarrow 0$ 
4:   Generate  $A_{aisle,R} \leftarrow \text{Split}(A_R, \frac{1}{n_a})$ 
5:   repeat
6:      $n_{it} \leftarrow n_{it} + 1$ 
7:     if  $n_{it} = 1$  then
8:        $A_{aisle}^{(n_{it})} \leftarrow A_{aisle,R}$ 
9:     else if  $n_{it} > 1$  then
10:      Generate  $A_S^{(n_{it})} \leftarrow \text{Merge}(A_{S_{re-en}}^{(n_{it}-1)}, A_{S_{rep}})$ 
11:      Generate  $A_{aisle,S}^{(n_{it})} \leftarrow \text{Split}(A_S^{(n_{it})}, \frac{1}{n_a})$ 
12:      Generate  $A_{aisle}^{(n_{it})} \leftarrow \text{Merge}(A_{aisle,S}^{(n_{it})}, A_{aisle,R})$ 
13:    end if
14:    Generate  $W_{aisle}^{(n_{it})} \leftarrow G|G|1(A_{aisle}^{(n_{it})}, B_{aisle})$ 
15:    Generate  $D_{aisle}^{(n_{it})} \leftarrow G|G|1(A_{aisle}^{(n_{it})}, B_{aisle})$ 
16:    Generate  $D_{aisle,R}^{(n_{it})} \leftarrow \text{Split}(D_{aisle}^{(n_{it})}, p_R)$ 
```

---

```

17:       $j \leftarrow 1$ 
18:      Generate  $D_{out,j}^{(n_{it})} \leftarrow \text{Merge}(D_{aisle,R}^{(n_{it})}, n_a\text{-times})$ 
19:      if  $n_{ps} > 0$  then
20:          repeat
21:              Generate  $A_{ps,j}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, \frac{p_{ps}}{n_{ps}})$ 
22:              Generate  $A_{\bar{ps},j}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, 1 - \frac{p_{ps}}{n_{ps}})$ 
23:              Generate  $D_{ps,j}^{(n_{it})} \leftarrow G|G|1(A_{ps,j}^{(n_{it})}, B_{ps})$ 
24:              Generate  $D_{out,j+1}^{(n_{it})} \leftarrow \text{Merge}(A_{\bar{ps},j}^{(n_{it})}, D_{ps,j}^{(n_{it})})$ 
25:               $j \leftarrow j + 1$ 
26:          until  $j > n_{ps}$ 
27:      end if
28:      Generate  $D_{out}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, 1 - p_{ps} + p_{ps} \cdot p_{em})$ 
29:      Generate  $A_{S_{re-en}}^{(n_{it})} \leftarrow \text{Split}(D_{out,j}^{(n_{it})}, p_{ps} - p_{ps} \cdot p_{em})$ 
30:       $T_R^{(n_{it})} \leftarrow W_{aisle}^{(n_{it})} \otimes B_{aisle,R}$ 
31:      until  $|E[T_R^{(n_{it})}] - E[T_R^{(n_{it}-1)}]| < \varepsilon_{it}$ 
32:       $T_R \leftarrow T_R^{(n_{it})}$ 
33:       $D_{out} \leftarrow D_{out}^{(n_{it})}$ 
34:      Generate  $Q_{aisle} \leftarrow G|G|1(A_{aisle}^{(n_{it})}, B_{aisle})$ 
35:      Generate  $Q_{aisle,S} \leftarrow \text{Binomial}(Q_{aisle}, 1 - p_R)$ 
36:      if  $n_{ps} > 0$  then
37:           $j \leftarrow 1$ 
38:          repeat
39:              Generate  $Q_{ps,j} \leftarrow G|G|1(A_{ps,j}^{(n_{it})}, B_{ps})$ 
40:               $j \leftarrow j + 1$ 
41:          until  $j > n_{ps}$ 
42:      end if
43:      return  $T_R, D_{out}, Q_{aisle,S}, Q_{ps,1}, \dots, Q_{ps,n_{ps}}$ 
44: end procedure

```

---



# **6 Validation and numerical evaluation**

In this chapter, we evaluate the approximation quality of the decomposition approach by comparing its results to the results obtained by a discrete-event simulation. Furthermore, we show how to use the discrete-time models to obtain cost efficient system configurations in an application example. Within the example, we conduct a numerical evaluation to investigate the influence of various parameters such as multi-level shuttles, re-entrant bins, and the reduction of the variability of the picking time as well as the storage and retrieval transaction inter-arrival times.

## **6.1 Validation**

In the absence of exact analytical results, we determine the deviations of the performance measures obtained by our decomposition approach to the values obtained by discrete-event simulation for a large number of different system configurations. The simulation software used is AnyLogic, version 7.3.1. Before we present the results of the validation, we first describe the sequence of events during the simulation.

### **6.1.1 Description of the discrete-event simulation**

The main assumptions of the simulation are identical to the decomposition approach, i.e. we assume the same physical design of the SBS/RS, the same resource characteristics and the same control policies. Moreover, as in the decomposition approach, we assume that the physical buffer capacities of the system are designed large enough such that there is no blocking. Due to the

random storage assignment rule and the equal access frequency, we do not keep track of the actual fill rate of the storage rack, i.e. we do not record the empty/occupied storage locations during the simulation. In contrast to the decomposition approach, the simulation treats each arriving transaction as a single customer class with deterministic routing to the target destination and state dependent service times when using the shuttles or lifts. It means that the time span in which a transaction uses a shuttle or a lift depends on the actual traveled distances, the velocity and acceleration/deceleration rates of the vehicles and the lifts as well as the transfer times. Furthermore, continuous inter-arrival time distributions of incoming retrieval transactions and storage transactions for replenishment are not discretized to the next time increment. The service times, when using the resources, are also not discretized to the next time increment. Given the average velocity of the roller conveyors outside of the SBS/RS, the times on different parts of the conveyor system depend on the length of these parts as well as the time on the merging elements. We assume a continuous split and merge of bins on the conveyor. Whereas the split does not delay the flow of bins, i.e. it does not lead to queues, the FCFS merge of bins may lead to waiting times. We assume that the service time of the bin on the merging element is larger than the required minimum inter-arrival time between two bins. In the following, for both the tier-captive and the tier-to-tier configurations we describe the sequence of events that take place during the retrieval or storage of a bin.

## Retrieval of a bin

1. Create the retrieval transaction based on the i.i.d. random variable  $A_R$ .
2. Randomly assign the storage location from where the bin has to be retrieved (aisle, tier, column).
3. Given  $p_{ps}$ , randomly decide whether the bin is routed to a picking station or out of the system.
4. If the bin is routed to a picking station:
  - Given  $n_{ps}$ , randomly assign a picking station.
  - Given  $p_{em}$ , randomly decide whether the bin is empty after picking.
5. In a tier-captive configuration:

- a) The transaction requests the vehicle.
  - b) The transaction waits for the vehicle (FCFS).
  - c) The vehicle travels from its idle position to the retrieval location, loads the bin, travels with the bin to the unloading point of the tier and unloads the bin.
  - d) The transaction requests the outgoing lift.
  - e) The transaction waits for the outgoing lift (FCFS).
  - f) The lift travels from its idle position to the retrieval tier, loads the bin, travels with the bin to the output point of the aisle and unloads the bin.
6. In a tier-to-tier configuration:
- a) The transaction requests the vehicle.
  - b) The transaction waits for the vehicle (FCFS).
  - c) If the vehicle is located on the same tier as the retrieval location, the vehicle travels from its idle position to the retrieval position, loads the bin and travels with the bin onto the lift at the (un)loading point of the tier. Afterward, the lift travels with the vehicle from the retrieval tier to the output point of the aisle, where the vehicle unloads the bin.
  - d) If the vehicle is located on another tier than the retrieval position, the vehicle travels from its idle position onto the lift at the (un)loading point of the tier and uses the lift to travel to the retrieval tier. At the retrieval tier, it travels to the retrieval location, loads the bin and travels with the bin back onto the lift at the (un)loading point of the tier. Afterward, the lift travels with the vehicle from the retrieval tier to the output point of the aisle, where the vehicle unloads the bin.
  - e) If the vehicle is located on the lift at the output point of the aisle, the vehicle directly uses the lift to travel to the retrieval tier. At the retrieval tier, it travels to the retrieval position, loads the bin and travels with the bin back onto the lift at the (un)loading point of the tier. Afterward, the lift travels with the vehicle from the retrieval tier to the output point of the aisle, where the vehicle unloads the bin.

7. The bin travels to its destination using the shortest route on the conveyor system. At merging elements, the bin first requests the element, then waits for the element to be empty (FCFS), and finally passes the element.
  - If the bin is routed to a picking station, the bin first waits for the picker (FCFS). Afterward, it is used by the picker according to the i.i.d. random variable  $B_{ps}$ .
  - If the bin is not routed to a picking station, the bin leaves the system.
8. If the bin was used by a picker and subsequently is empty, the bin is routed out of the system.
9. If the bin was used by a picker and subsequently is not empty, create a re-entering storage transaction and randomly assign a storage location where the bin has to be stored (aisle, tier, column). Afterward, the bin is routed on the conveyor system back to the SBS/RS (see below).

### Storage of a bin

1. In case of bins for replenishment, create the storage transaction based on the i.i.d. random variable  $A_{S_{rep}}$  and randomly assign the storage location where the bin has to be stored (aisle, tier, column).
2. In case of re-entering bins after picking (see above), create the storage transaction at the picking station and randomly assign a storage location (aisle, tier, column).
3. The bin travels to the incoming lift of its target aisle using the shortest route on the conveyor system. At merging elements, the bin first requests the element, then waits for the element to be empty (FCFS), and finally passes the element.
4. In a tier-captive configuration:
  - a) The transaction requests the incoming lift.
  - b) The transaction waits for the incoming lift (FCFS).
  - c) The lift travels from its idle position to the input point of the aisle, loads the bin, travels with the bin to the target tier and unloads the bin.
  - d) The transaction requests the vehicle.

- e) The transaction waits for the vehicle (FCFS).
  - f) The vehicle travels from its idle position to the loading point of the tier, loads the bin, travels with the bin to the storage location and unloads the bin.
5. In a tier-to-tier configuration:
- a) The transaction requests the vehicle.
  - b) The transaction waits for the vehicle (FCFS).
  - c) If the vehicle is located on a tier inside the SBS/RS, the vehicle travels from its idle position onto the lift at the (un)loading point of the tier. Afterward, it uses the lift to travel to the input point of the aisle, where it loads the bin. It also uses the lift to travel with the bin to the target tier. At the target tier, it travels to the storage location and unloads the bin.
  - d) If the vehicle is located on the lift at the output point of the aisle, the vehicle directly uses the lift to travel to the input point of the aisle, where it loads the bin. It also uses the lift to travel with the bin to the target tier. At the target tier, it travels to the storage position and unloads the bin.

Given those sequences, for every simulation run we are able to compute the performance measures of the resources such as the average vehicle utilization. Furthermore, we are able to determine the performance measures for each transaction such as the retrieval transaction time. Given all retrieval transaction times of a simulation run, we can calculate their average value and the quantiles of the resulting distribution.

### 6.1.2 Evaluation of the approximation quality

In order to evaluate the approximation quality, we compare the performance measures obtained by our approach, such as the expected value and the 95% quantile of the retrieval transaction time, to the ensemble averages of 10 independent simulation replications. In the decomposition approach, we stop the iterations if  $|E[T_R^{(n_u)}] - E[T_R^{(n_u-1)}]| < 0.001$ . In each replication of the simulation, we stop the simulation after 10,000,000 transactions are fulfilled and leave the system. We start collecting data after a warm up period of 10,000 fulfilled

transactions. Previous simulation experiments have shown that these numbers are needed to reach the steady-state. Afterward, we compute the absolute and relative deviations of the key values of the performance measures obtained by the decomposition approach ( $\text{value}^{DA}$ ) to the ensemble averages of the discrete-event simulation ( $\text{value}^{DES}$ ). The relative deviations are of interest for performance measures that can assume large values such as the retrieval transaction time, whereas the absolute deviations are of interest for performance measures that only assume small values and, therefore, can be easily interpreted directly (e.g., the number of waiting bins at the arrival instant). For a given system configuration  $c$ , we compute the deviations  $\Delta_c^{abs}$  and  $\Delta_c^{rel}$  as follows.

$$\Delta_c^{abs} = \text{value}_c^{DA} - \text{value}_c^{DES} \quad (6.1)$$

$$\Delta_c^{rel} = \frac{\text{value}_c^{DA} - \text{value}_c^{DES}}{\text{value}_c^{DES}} \quad (6.2)$$

Given those deviations, we are able to determine for the  $n_{config}^{val}$  analyzed configurations the cumulative distribution function of the deviations. This gives us an understanding about the approximation quality of the decomposition approach. In addition to the CDF, we are interested in the average absolute deviations  $|\Delta^{abs}|_{avg}$  and  $|\Delta^{rel}|_{avg}$ , which we compute as follows.

$$|\Delta^{abs}|_{avg} = \frac{1}{n_{config}^{val}} \cdot \sum_{c=1}^{n_{config}^{val}} |\text{value}_c^{DA} - \text{value}_c^{DES}| \quad (6.3)$$

$$|\Delta^{rel}|_{avg} = \frac{1}{n_{config}^{val}} \cdot \sum_{c=1}^{n_{config}^{val}} \left| \frac{\text{value}_c^{DA} - \text{value}_c^{DES}}{\text{value}_c^{DES}} \right| \quad (6.4)$$

As stated in the previous chapters, we are mainly interested in the retrieval transaction time distribution and the distributions of the number of bins waiting in front of the incoming lifts as well as the picking stations. Given the varying number of incoming lift and picking stations, we focus on the queue length distribution of the first incoming lift station, which in the decomposition approach is identical for all  $\text{lift}_{in}$  stations due to the random storage assignment

rule, and the queue length distribution of the first picking station. Furthermore, the inter-departure time distribution of the transactions that are leaving the system might be of interest if we use it as input for subsequent discrete-time building blocks. Since we are not just interested in the expected values of the performance measures, but also in their quantiles, we determine the deviations for both the expected values and the 95% quantiles of the following performance measures.

- $T_R$ : retrieval transaction time
- $D_{out}$ : inter-departure time of the transactions leaving the system
- $Q_{lift_{in}}$ : queue length of the first lift<sub>in</sub> station at the arrival instant (in case of a tier-captive system)
- $Q_{aisle,S}$ : storage transaction queue length of the first aisle station at the arrival instant (in case of a tier-to-tier system)
- $Q_{ps,1}$ : queue length of the first picking station at the arrival instant

Table 6.1 depicts the different configurations chosen to reflect a broad variety of different settings. To be close to real applications, we evaluate the approximation quality for configurations with an aisle length of up to  $100m$  and a height of up to  $8.64m$ . Moreover, we consider systems with up to 8 aisles. The data regarding the size of the equally sized storage columns as well as the velocities, acceleration/deceleration rates and transfer times of the lifts and the vehicles, which may have a maximum of 4 levels, are provided to us by a European material handling provider. Furthermore, we use retrieval inter-arrival times that are distributed according to an Exponential distribution or randomly generated discrete distributions with an scv of 0.25. Since the tier-captive configurations lead to a higher throughput than the tier-to-tier configurations, we assume different expected values of the retrieval inter-arrival time distributions. The inter-arrival times of the storage transactions for replenishment are distributed according to a Gamma distribution with an scv of 0.25. Due to the required minimum distance between two bins on the conveyor of  $0.7m$ , the Gamma distribution has a lower bound equal to  $0.35s$ . We want to investigate systems, in which the arrival rate of the storage transactions equals the arrival rate of the retrieval transactions. Therefore, we adjust the expected value of the storage transaction inter-arrival time distribution for replenishment based on the values of  $E[A_R]$ ,  $p_{ps}$  and  $p_{em}$ .

Table 6.1: Tested parameter configurations of the tier-captive and tier-to-tier SBS/RSS

Parameter	Values
$n_a$	{2, 4, 6, 8}
$n_l$	{1, 2, 4}
$n_t$	$\left\{ \frac{8}{n_l}, \frac{16}{n_l}, \frac{24}{n_l} \right\}$
$n_c$	{50, 100, 150, 200}
$p_{ps}$	{0.0, 0.5, 1.0}
$A_R$	{Exponential, Discrete( $cv^2 = 0.25$ )}
$E[A_R]$ (tier-captive)	{4.0s, 5.0s, 6.0s, 7.0s, 8.0s, 9.0s}
$E[A_R]$ (tier-to-tier)	{15s, 20s, 25s, 30s}
$B_{ps}$	Gamma with $cv^2 = 0.25$
$E[B_{ps}]$	10s
$n_{ps}$	$\left\lceil \frac{E[B_{ps}] \cdot p_{ps}}{0.9 \cdot E[A_R]} \right\rceil$
$p_{em}$	0.1
$A_{S_{rep}}$	Gamma with $cv^2 = 0.25$
$E[A_{S_{rep}}]$	$\frac{E[A_R]}{1 - p_{ps} + p_{ps} \cdot p_{em}}$
$d_c$	0.5m
$d_l$	0.36m
$d_{l_{in}}$	0.7m
$d_{l_{out}}$	1.5m
$v_l = v_{l_{in}} = v_{l_{out}}$	$5 \frac{m}{s}$
$a_l = a_{l_{in}} = a_{l_{out}}$	$5 \frac{m}{s^2}$
$t_{l_{in},trans} = t_{l_{out},trans}$	2.5s
$v_{v,x}$	$2 \frac{m}{s}$
$a_{v,x}$	$1 \frac{m}{s^2}$
$v_{v,y}$	$1 \frac{m}{s}$
$a_{v,y}$	$2 \frac{m}{s^2}$
$t_{v,trans}$	2.5s

The values of  $p_{ps}$  are set to 0.0, 0.5, or 1.0, whereas the value of  $p_{em}$  is equal to 0.1 (if there is picking). The service times at the picking stations are also distributed according to a Gamma distribution with an scv of 0.25 and a lower bound of 0.35s. For every possible configuration, we choose the required num-

ber of picking stations such that the utilization of a picking station is less than 90%. Therefore, the number of picking stations depends on the retrieval transaction stream that is routed to the picking stations and the service times of the picking stations, i.e.  $n_{ps} = \left\lceil \frac{E[B_{ps}] \cdot p_{ps}}{0.9 E[A_R]} \right\rceil$ .

In the decomposition approach, the Exponential distributions are approximated by Geometric distributions with a lower support of  $1 \cdot t_{inc}$ . Regarding the discretized Gamma distributions, the probability that the random variable assumes  $i \cdot t_{inc}$  is equal to the probability mass of the respective Gamma distribution between  $i \cdot t_{inc} - 0.5 \cdot t_{inc}$  and  $i \cdot t_{inc} + 0.5 \cdot t_{inc}$ . The probability mass between 0 and  $0.5 \cdot t_{inc}$  is added to  $1 \cdot t_{inc}$ . Both discretized distributions are truncated such that the cumulated probability of values beyond the upper support is smaller than 0.0001%. Afterward, the defective distributions are normalized.

The discrete inter-arrival distributions are randomly created. If they are distributed, for example, according to the random variable  $X$ , their lower and upper supports are  $1 \cdot t_{inc}$  and  $\lceil 2 \cdot E[X] \rceil$ , respectively. In the decomposition approach, the constant time increment  $t_{inc}$  is set to 1s. In the simulation, we assume that the velocity of the conveyor system is equal to  $2 \frac{m}{s}$  and the time that a bin needs to pass a merging element is equal to 0.5s. The discrete distributions as well as the distances of the different parts of the conveyor system are given in appendix A.

All possible combinations lead to 5,184 different configurations for the tier-captive system and 3,456 configurations for the tier-to-tier system. Out of those, we only consider the cases in which the maximum utilization of the resources inside the SBS/RS is between 50% and 90%. This leads to  $n_{config}^{val} = 1,098$  analyzed cases for the tier-captive system and  $n_{config}^{val} = 1,086$  analyzed cases for the tier-to-tier system. To show the impact of using the fast split approximation described in section 4.2.4 on the approximation quality and the computation times, we perform the computations with the decomposition approach for all configurations two times. In the first time, the split of a stochastic stream is done using the stochastic split method presented in section 4.2.3. The second time, we use for every split operation the fast split approximation (see section 4.2.4). The threshold value  $n_{FSA}$  is set equal to 12. Afterward, we compute the absolute and relative differences in the computation time for all configurations. The computations are performed on a Intel Xeon CPU E5-2630 v3 @ 2.40GHz with 2 processors and 64 GB RAM.

### Approximation quality of the tier-captive configurations

In general, the decomposition approach reaches a high approximation quality in the analyzed tier-captive configurations. Regarding the retrieval transaction time, the average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 1.41% and 1.91% when using the stochastic split method. Given the fast split approximation, the discrete-time approach reaches a high approximation quality, too. The average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 2.52% and 3.09%, respectively.

The cumulative distribution functions of the relative deviations of the retrieval transaction time  $T_R$  regarding the expected value  $E[T_R]$  and the 95% quantile  $T_{R,0.95}$  given the stochastic split and the fast split methods are displayed in figure 6.1.

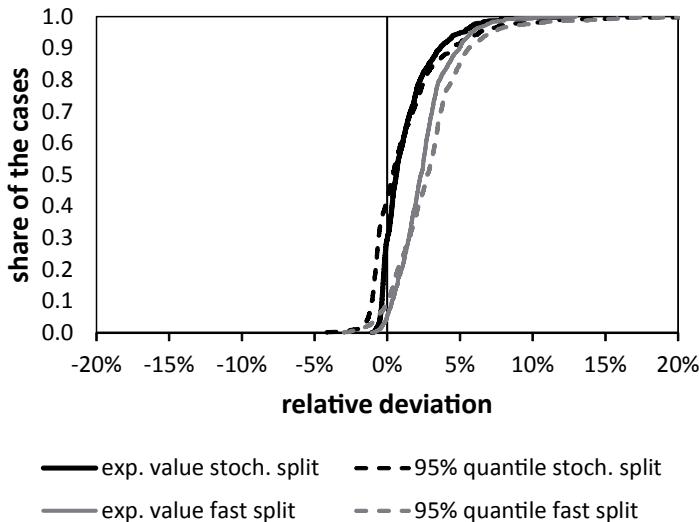


Figure 6.1: Cumulated distribution of the relative deviations of the retrieval transaction time  $T_R$  regarding the expected value  $E[T_R]$  and the 95% quantile  $T_{R,0.95}$  given the stochastic split and the fast split methods

It illustrates that most of the deviations are positive, meaning that the discrete-time approximation overestimates the performance measures in many cases. Moreover, it shows that in more than 90% of the cases, the relative deviations of both the expected value and the 95% quantile are less than 6%. An additional analysis reveals that the approximation reaches a very high quality if the arrival process of the retrieval transactions is a Poisson process. In these cases, the average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 0.38% and 0.81% when using the stochastic split method, and 1.30% and 1.39% when using the fast split approximation. The decomposition approach reaches a lower approximation quality when the inter-arrival times of the retrieval transactions are distributed according to discrete distribution with an scv of 0.25. In these cases, the average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 2.44% and 3.00% when using the stochastic split method and 3.75% and 4.79% when using the fast split approximation. Given a non-Poisson arrival process of the retrieval transactions, the decomposition approach overestimates the expected value of the retrieval transaction time distribution in all cases. Regarding the 95% quantile, it overestimates the performance measure in 98.72% of the cases given the stochastic split method and in all cases given the fast split approximation.

Regarding the inter-departure time of the transactions leaving the system, as expected, the deviations are close to zero in terms of the expected value of  $D_{out}$ . In terms of the 95% quantile  $D_{out,0.95}$ , the average absolute deviations  $|\Delta^{rel}|_{avg}$  equal 2.00% when using the stochastic split method and 3.40% given the fast split approximation. Again, the approximation quality of the 95% quantile is higher given a Poisson arrival process of the retrieval transactions (1.38% for the stochastic split and 2.77% for the fast split) and lower in case of an arrival process with an scv of 0.25 (2.61% for the stochastic split and 4.02% for the fast split). The cumulative distribution functions of the relative deviations of  $D_{out}$  regarding the expected value  $E[D_{out}]$  and the 95% quantile  $D_{out,0.95}$  given the stochastic split and the fast split methods are displayed in figure 6.2. Since the deviations regarding the expected value  $E[D_{out}]$  are close to zero for both split methods, the graphs overlap each other such that only the graph in gray color can be seen. The same applies for the negative deviations of the 95% quantile. As for the retrieval transaction time, over 90% of the deviations are below 6%.

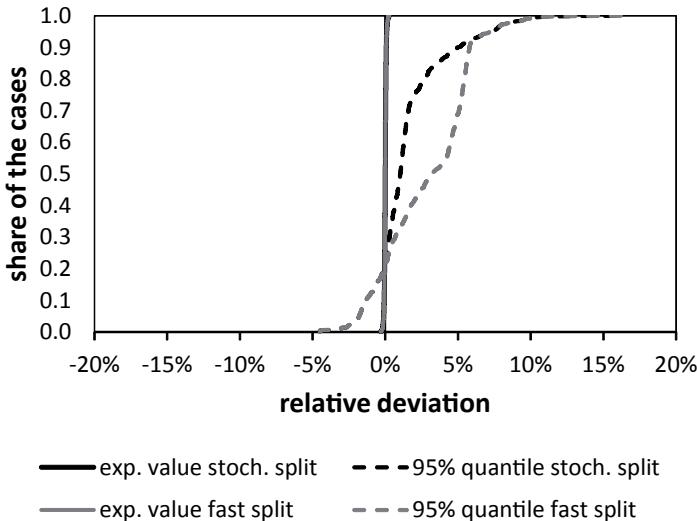


Figure 6.2: Cumulated distribution of the relative deviations of the inter-departure time of the transactions leaving the system  $D_{out}$  regarding the expected value  $E[D_{out}]$  and the 95% quantile  $D_{out,0.95}$  given the stochastic split and the fast split methods

Since we are interested in the absolute deviations regarding the number of bins waiting in front of an incoming lift station at the arrival instant, we compute the average absolute deviations  $|\Delta^{abs}|_{avg}$  of the expected value and the 95% quantile of  $Q_{lift_{in}}$ . Regardless of the used method to split the stochastic streams, the average deviations for the expected value and the 95% quantile equal 0.05 and 0.26 bins, respectively. Moreover, in over 99% of the analyzed configurations, the absolute deviation was less than or equal to 1 bin, and the largest deviation was 2 bins. This is also evident from the cumulative distribution functions of the absolute deviations, which are displayed in figure 6.3. Since the deviations are very similar for both split methods, the graphs overlap each other such that only the graphs in gray color can be seen.

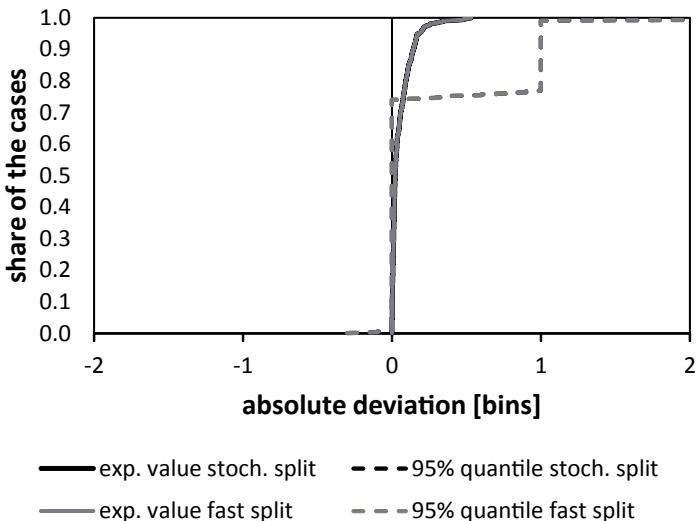


Figure 6.3: Cumulated distribution of the absolute deviations of the number of bins waiting in front of an incoming lift station at the arrival instant  $Q_{lift_{in}}$  regarding the expected value  $E[Q_{lift_{in}}]$  and the 95% quantile  $Q_{lift_{in},0.95}$  given the stochastic split and the fast split methods

Regarding the absolute deviations of the number of bins waiting in front of the first picking station, the approximation quality is slightly worse than in the case of the number of bins waiting in front of the first incoming lift station. The average absolute deviations  $|\Delta^{abs}|_{avg}$  of the expected value and the 95% quantile of  $Q_{ps,1}$  are equal to 0.15 and 0.55 bins, respectively (regardless of the used method to split the stochastic streams). The cumulative distribution functions of the absolute deviations, which are displayed in figure 6.4, show that in almost 90% of the analyzed configurations the absolute deviation was less than or equal to 1 bin, and the largest deviation was 3 bins (for the 95% quantile). Since the deviations are very similar for both split methods, the graphs overlap each other such that only the graphs in gray color can be seen. As before, the approximation quality is higher given a Poisson arrival process of the retrieval transactions. In these cases, the average absolute deviations  $|\Delta^{abs}|_{avg}$  of the

expected value and the 95% quantile equal 0.08 and 0.28 bins, regardless of the used method to split the stochastic streams. On the contrary, the decomposition approach reaches a lower approximation quality given a non-Poisson arrival process of the retrieval transactions. In these cases, the average absolute deviations  $|\Delta^{abs}|_{avg}$  of the expected value and the 95% quantile equal 0.22 and 0.81 bins, regardless of the method used to split the stochastic streams. An overview of the average absolute deviations  $|\Delta^{rel}|_{avg}$  and  $|\Delta^{abs}|_{avg}$  of the performance measures is given in tables 6.2 and 6.3, respectively.

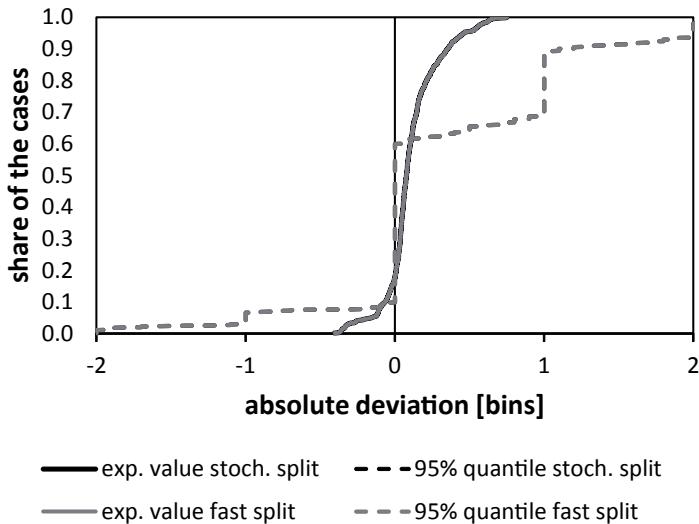


Figure 6.4: Cumulated distribution of the absolute deviations of the number of bins waiting in front of the first picking station  $Q_{ps,1}$  regarding the expected value  $E[Q_{ps,1}]$  and the 95% quantile  $Q_{ps,1,0.95}$  given the stochastic split and the fast split methods

Table 6.2: Overview of the average absolute deviations  $|\Delta^{rel}|_{avg}$  of  $E[T_R]$ ,  $T_{R,0.95}$ ,  $E[D_{out}]$ , and  $D_{out,0.95}$

split operation	perform. measure	$ \Delta^{rel} _{avg}^*$	$ \Delta^{rel} _{avg}^{**}$	$ \Delta^{rel} _{avg}^{***}$
stoch. split	$E[T_R]$	1.41%	0.38%	2.44%
	$T_{R,0.95}$	1.91%	0.81%	3.00%
	$E[D_{out}]$	0.05%	0.05%	0.04%
	$D_{out,0.95}$	2.00%	1.38%	2.61%
fast split	$E[T_R]$	2.52%	1.30%	3.75%
	$T_{R,0.95}$	3.09%	1.39%	4.79%
	$E[D_{out}]$	0.05%	0.05%	0.04%
	$D_{out,0.95}$	3.40%	2.77%	4.02%

\* All configurations  $n_{config}^{val}$

\*\* $A_R$ : Poisson process

\*\*\* $A_R$ : non-Poisson process

Table 6.3: Overview of the average absolute deviations  $|\Delta^{abs}|_{avg}$  of  $E[Q_{liftin}]$ ,  $Q_{liftin,0.95}$ ,  $E[Q_{ps,1}]$ , and  $Q_{ps,1,0.95}$

split operation	perform. measure	$ \Delta^{abs} _{avg}^*$ [bins]	$ \Delta^{abs} _{avg}^{**}$ [bins]	$ \Delta^{abs} _{avg}^{***}$ [bins]
stoch. split	$E[Q_{liftin}]$	0.05	0.05	0.06
	$Q_{liftin,0.95}$	0.26	0.23	0.28
	$E[Q_{ps,1}]$	0.15	0.08	0.22
	$Q_{ps,1,0.95}$	0.55	0.28	0.81
fast split	$E[Q_{liftin}]$	0.05	0.05	0.06
	$Q_{liftin,0.95}$	0.26	0.23	0.28
	$E[Q_{ps,1}]$	0.15	0.08	0.22
	$Q_{ps,1,0.95}$	0.55	0.28	0.81

\* All configurations  $n_{config}^{val}$

\*\* $A_R$ : Poisson process

\*\*\* $A_R$ : non-Poisson process

To demonstrate the accuracy of the decomposition approach, we present in figure 6.5 the retrieval transaction time distributions that are obtained by the decomposition approach using the stochastic split method and the discrete-event simulation for a configuration with the following parameters: 4 aisles, 8 storage locations in vertical direction and 50 locations in horizontal direction, 2 level per tier, a pick probability of  $p_{ps} = 0.5$ , and an expected value of 4.0s for the retrieval transaction inter-arrival time, which is distributed according to a discrete distribution with an scv of 0.25. For a better visual comparison, the retrieval transaction times obtained in the simulation are rounded to the next time increment, i.e. to the next second. Afterward, for every time increment the average probability of 10 simulation runs is computed.

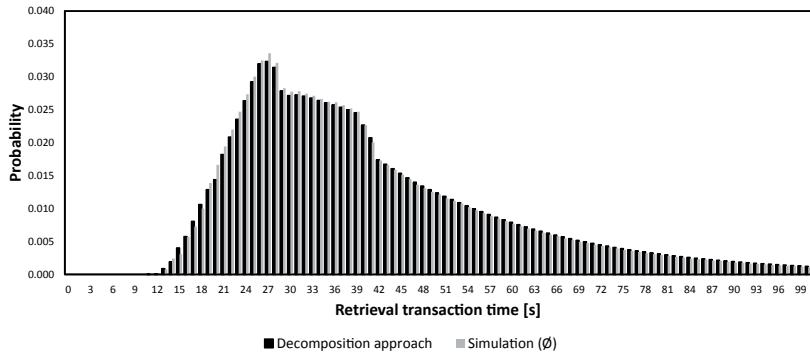


Figure 6.5: Comparison of the probability distributions of the retrieval transaction time  $T_R$  obtained by the decomposition approach using the stochastic split method and the ensemble average ( $\bar{\emptyset}$ ) of 10 simulation runs for one of the tier-captive configurations

In this configuration, the decomposition approach reaches an approximation quality, in which the deviations are similar to the average deviations of all analyzed configurations. The expected value of the retrieval transaction time of the decomposition approach is 42.88s, thus 1.52% higher than the expected value of the discrete-event simulation, which is 42.23s. Moreover, the 95% quantile of the decomposition approach is 86s, thus 1.98% higher than the ensemble average of the 10 simulation runs, which equals 84.33s. We see

that the shapes of the distributions are very alike, meaning that also the other quantiles of the distribution are met quite well.

The 95% confidence intervals of the performance measures of all simulation runs are computed as well. Regarding the expected value and the 95% quantile of the retrieval transaction time, the average relative confidence interval is equal to  $\pm 0.15\%$  and  $\pm 0.27\%$ , respectively. In terms of the expected value and the 95% quantile of the inter-departure time of the transactions leaving the system, the average relative confidence interval is equal to  $\pm 0.14\%$  and  $\pm 0.22\%$ , respectively. For the expected value and the 95% quantile of the number of bins waiting in front of the incoming lift stations and the first picking station at the arrival instant, we receive average absolute confidence intervals of  $\pm 0.001$  bins ( $E[Q_{liftin}]$ ),  $\pm 0.010$  bins ( $Q_{liftin,0.95}$ ),  $\pm 0.011$  bins ( $E[Q_{ps,1}]$ ), and  $\pm 0.058$  bins ( $Q_{ps,1,0.95}$ ), respectively.

In terms of the computation times, the average computation time of a configuration decreases by 74.32% from 36.10s to 9.27s when using the fast split approximation instead of the stochastic split method. It is obvious from the shown figures that the fast split approximation leads to a lower approximation quality in case of the retrieval transaction time distribution and the inter-departure time distribution of the transactions leaving the system. However, the differences are rather small in terms of the absolute percentage differences. Therefore, the fast split approximation might be of interest during the early planning phase of such systems given a large number of possible system configurations that have to be compared to each other.

### **Approximation quality of the tier-to-tier configurations**

In general, the decomposition approach reaches a high approximation quality. However, the approximation quality is lower than the one reached for the tier-captive configurations. In the following, we will present the deviations of the decomposition approach using the stochastic split method. The values obtained by using the fast split approximation are given in brackets.

Regarding the retrieval transaction time, the average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 2.57% (2.61%) and 4.12% (4.19%), respectively. The cumulative distribution functions of the relative deviations of the retrieval transaction time  $T_R$  are displayed in figure 6.6.

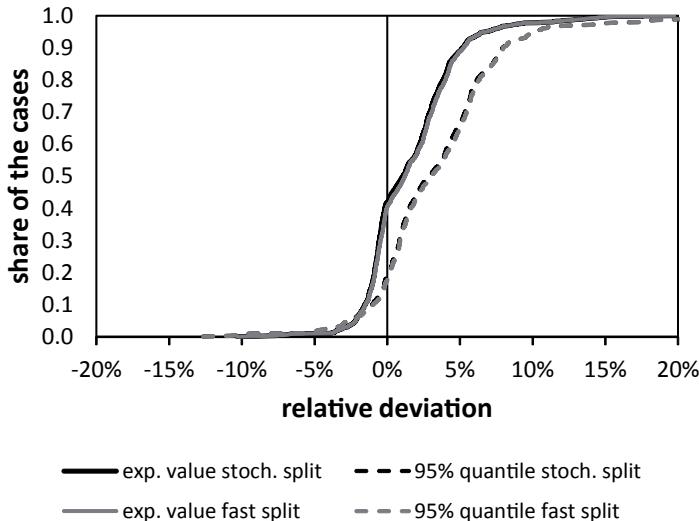


Figure 6.6: Cumulated distribution of the relative deviations of the retrieval transaction time  $T_R$  regarding the expected value  $E[T_R]$  and the 95% quantile  $T_{R,0.95}$  given the stochastic split and the fast split methods

It shows that the deviations are both negative and positive. Furthermore, it shows that in over 90% of the cases the relative deviations of both the expected value and the 95% quantile are less than 10%. Again, since the deviations are very similar for both split methods, the graphs overlap each other such that only the graphs in gray color can be seen. An additional analysis reveals that the decomposition approach reaches a high approximation quality if the arrival process of the retrieval transactions is a Poisson process and the number of aisles is large. In the cases with Poisson arrivals and 8 aisles, the average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 0.50% (0.48%) and 1.11% (1.25%), respectively. On the contrary, the decomposition approach reaches a low approximation quality for two aisles and a non-Poisson arrival process of the retrieval transactions. In these cases, the average absolute deviations  $|\Delta^{rel}|_{avg}$  of the expected value and the 95% quantile equal 12.27% (12.27%) and 18.51% (18.51%), respectively. Given a small

number of aisles and a non-Poisson arrival process of the retrieval transactions, the decomposition approach overestimates the performance measures of the retrieval transaction time distribution in all cases.

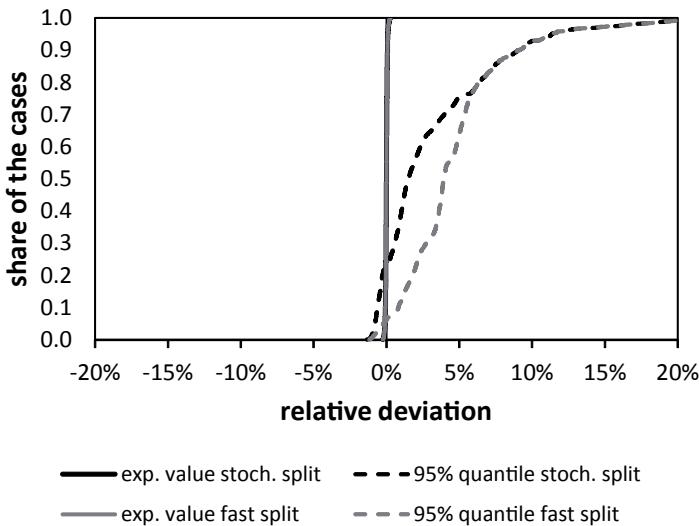


Figure 6.7: Cumulated distribution of the relative deviations of the inter-departure time of the transactions leaving the system  $D_{out}$  regarding the expected value  $E[D_{out}]$  and the 95% quantile  $D_{out,0.95}$  given the stochastic split and the fast split methods

Regarding the inter-departure time of the transactions leaving the system, the deviations are close to zero in terms of the expected value of  $D_{out}$ . In terms of the 95% quantile  $D_{out,0.95}$ , the average absolute deviations  $|\Delta^{rel}|_{avg}$  equal 3.45% (4.70%). Again, the approximation quality of the 95% quantile is higher given a Poisson arrival process of the retrieval transactions and 8 aisles (1.17% for the stochastic split and 2.20% for the fast split) and lower in case of two aisles and a retrieval arrival process with an scv of 0.25 (10.65% for the stochastic split and 12.11% for the fast split). The cumulative distribution functions of the relative deviations of  $D_{out}$  regarding the expected value  $E[D_{out}]$  and the 95% quantile  $D_{out,0.95}$  given the stochastic split and the fast split approximation are displayed in figure 6.7. Since the deviations regarding

the expected value  $E[D_{out}]$  are close to zero for both split methods, the graphs overlap each other such that only the graph in gray color can be seen. As for the retrieval transaction time, over 90% of the deviations are below 10%.

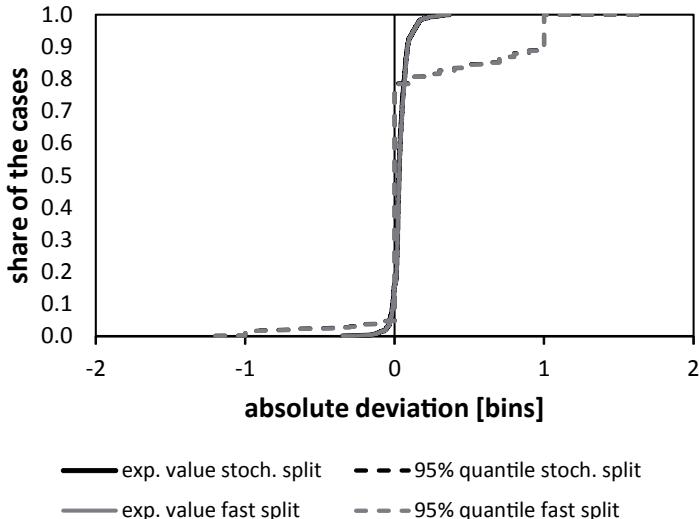


Figure 6.8: Cumulated distribution of the absolute deviations of the number of bins waiting in front of an aisle station at the arrival instant  $Q_{aisle,S}$  regarding the expected value  $E[Q_{aisle,S}]$  and the 95% quantile  $Q_{aisle,S,0.95}$  given the stochastic split and the fast split methods

Furthermore, we compute the average absolute deviations  $|\Delta^{abs}|_{avg}$  of the expected value and the 95% quantile of  $Q_{aisle,S}$ . In general, the average absolute deviations for the expected value and the 95% quantile equal 0.05 (0.05) and 0.19 (0.19) bins, respectively. Moreover, in over 99% of the analyzed configurations the absolute deviation is less than or equal to 1 bin. This is also evident from the cumulative distribution functions of the absolute deviations, which are displayed in figure 6.8. As before, since the deviations are very similar for both split methods, the graphs overlap each other such that only the graphs in gray color can be seen.

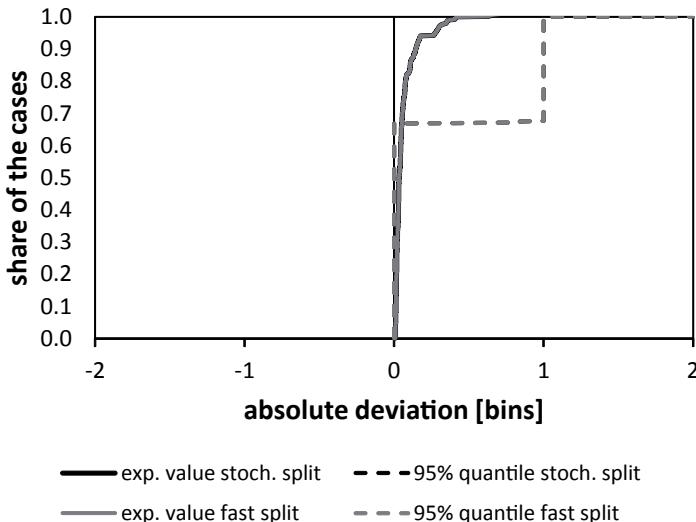


Figure 6.9: Cumulated distribution of the absolute deviations of the number of bins waiting in front of the first picking station  $Q_{ps,1}$  regarding the expected value  $E[Q_{ps,1}]$  and the 95% quantile  $Q_{ps,1,0.95}$  given the stochastic split and the fast split methods

As in the tier-captive configurations, the absolute deviations of the number of bins waiting in front of the first picking station are slightly higher than the absolute deviations of the number of bins waiting in front of the first aisle. The average absolute deviations  $|\Delta^{abs}|_{avg}$  of the expected value and the 95% quantile of  $Q_{ps,1}$  are equal to 0.06 (0.06) and 0.33 (0.33) bins, respectively. The cumulative distribution functions of the absolute deviations, which are displayed in figure 6.9, show that in almost 99% of the analyzed configurations the absolute deviation is less than or equal to 1 bin, and the largest deviation is 3 bins (for the 95% quantile). As before, since the deviations are very similar for both split methods, the graphs overlap each other such that only the graphs in gray color can be seen.

Again, the average computation time of a configuration decreases drastically when using the fast split approximation (it decreases by 50.79% from 189.35s to 93.18s). The differences in deviations between the two methods are rather small. Therefore, the fast split approximation might also be of interest during the early planning phase of tier-to-tier systems, in which a large range of different configurations must be analyzed. An overview of the average absolute deviations  $|\Delta^{rel}|_{avg}$  and  $|\Delta^{abs}|_{avg}$  of the performance measures is given in tables 6.4 and 6.5, respectively.

Table 6.4: Overview of the average absolute deviations  $|\Delta^{rel}|_{avg}$  of  $E[T_R]$ ,  $T_{R,0.95}$ ,  $E[D_{out}]$ , and  $D_{out,0.95}$

split operation	perform. measure	$ \Delta^{rel} _{avg}^*$	$ \Delta^{rel} _{avg}^{**}$	$ \Delta^{rel} _{avg}^{***}$
stoch. split	$E[T_R]$	2.57%	0.50%	12.27%
	$T_{R,0.95}$	4.12%	1.11%	18.51%
	$E[D_{out}]$	0.05%	0.05%	0.04%
	$D_{out,0.95}$	3.45%	1.17%	10.65%
fast split	$E[T_R]$	2.61%	0.48%	12.27%
	$T_{R,0.95}$	4.19%	1.25%	18.51%
	$E[D_{out}]$	0.05%	0.05%	0.04%
	$D_{out,0.95}$	4.70%	2.20%	12.11%

\* All configurations  $n_{config}^{val}$

\*\* $A_R$ : Poisson process and 8 aisles

\*\*\* $A_R$ : non-Poisson process and 2 aisles

The 95% confidence intervals of the performance measures of all simulation runs are computed as well. Regarding the expected value and the 95% quantile of the retrieval transaction time, the average relative confidence interval is equal to  $\pm 0.21\%$  and  $\pm 0.35\%$ , respectively. In terms of the expected value and the 95% quantile of the inter-departure time of the transactions leaving the system, the average relative confidence interval is equal to  $\pm 0.14\%$  and  $\pm 0.21\%$ , respectively. For the expected value and the 95% quantile of the number of bins waiting in front of the aisle stations and the first picking station at the ar-

rival instant, we receive average absolute confidence intervals of  $\pm 0.007$  bins ( $E[Q_{aisle,S}]$ ),  $\pm 0.042$  bins ( $Q_{aisle,S,0.95}$ ),  $\pm 0.001$  bins ( $E[Q_{ps,1}]$ ), and  $\pm 0.002$  bins ( $Q_{ps,1,0.95}$ ), respectively.

Table 6.5: Overview of the average absolute deviations  $|\Delta^{abs}|_{avg}$  of  $E[Q_{aisle,S}]$ ,  $Q_{aisle,S,0.95}$ ,  $E[Q_{ps,1}]$ , and  $Q_{ps,1,0.95}$

split operation	perform. measure	$ \Delta^{abs} _{avg}^*$ [bins]	$ \Delta^{abs} _{avg}^{**}$ [bins]	$ \Delta^{abs} _{avg}^{***}$ [bins]
stoch. split	$E[Q_{aisle,S}]$	0.05	0.02	0.18
	$Q_{aisle,S,0.95}$	0.19	0.16	0.68
	$E[Q_{ps,1}]$	0.06	0.05	0.02
	$Q_{ps,1,0.95}$	0.33	0.18	0.09
fast split	$E[Q_{aisle,S}]$	0.05	0.02	0.18
	$Q_{aisle,S,0.95}$	0.19	0.16	0.68
	$E[Q_{ps,1}]$	0.06	0.05	0.02
	$Q_{ps,1,0.95}$	0.33	0.18	0.09

\* All configurations  $n_{config}^{val}$

\*\* $A_R$ : Poisson process and 8 aisles

\*\*\* $A_R$ : non-Poisson process and 2 aisles

## 6.2 Numerical evaluation

In the following application example, we show how the decomposition approach can be used to determine system configurations that fulfill given requirements such as system capacity, throughput and service level in terms of the 95% quantile of the retrieval transaction time. Within the application example, we quantify the impact of multi-level shuttles, re-entrant bins, and the variability reduction of the retrieval inter-arrival time, the picking service times, and the inter-arrival time of the storage transactions for replenishment, on the design decision.

The system under investigation is a tier-captive SBS/RS as presented in chapter 5.1. The size of the storage columns as well as the velocities, acceleration/deceleration rates and transfer times of the vehicles and lifts are identical

to the ones used in the validation section. At first, we assume that the number of vehicle levels is 1 (i.e.  $n_l = 1$ ), the inter-arrival times of the retrieval transactions are exponentially distributed, and the inter-arrival times of the replenishment storage transactions as well as the service times of the picking stations are distributed according to a Gamma distribution with an scv of 0.25. We assume that the arrival rate of the storage transactions is equal to the arrival rate of the retrieval transactions, i.e. the value of  $E[A_{S_{rep}}]$  is computed in the same way as in the validation section. Furthermore, the expected value of the picking time is identical to the one in the validation section, i.e.  $E[B_{ps}] = 10s$ . The probability of a retrieval transaction being routed to a picking station is  $p_{ps} = 0.5$ , and the probability that a bin is empty after picking is  $p_{em} = 0.1$ . Again, the required number of picking stations is chosen such that the utilization of a picking station is less than 90%.

The chosen tier-captive SBS/RS design should fulfill the requirements given in table 6.6. The system should have a capacity of at least 20,000 storage locations. Moreover, the height of the system may be at most  $10m$ , the length at most  $100m$ , and the width at most  $10m$ . In addition, the design should fulfill the requirement that the maximum utilization of both the vehicles and the lifts is at most 90%, and that 95% of the retrieval transactions leave the SBS/RS within 120s, i.e. that the 95% quantile of the retrieval transaction time distribution is at most 120s.

Table 6.6: Requirements for the tier-captive SBS/RS

Parameter	Values
Storage capacity	$Cap = 20,000$
Maximum height the system	$max_H = 10m$
Maximum length of the system	$max_L = 100m$
Maximum width of the system	$max_W = 10m$
Throughput of the retrieval transactions	$\lambda_R = 1,000 \frac{1}{h}$
Maximum utilization	$\rho_{max} = 0.9$
Maximum value of the 95% quantile of the retrieval transaction time	$T_{R,0.95} = 120s$

In the following sections, we show how to efficiently design the tier-captive SBS/RS under the given requirements and quantify the impact of multi-level shuttles, re-entrant bins, and the variability of the random variables on the design decision. The same we perform for a tier-to-tier SBS/RS as presented in chapter 5.2 with the same assumptions and requirements. Due to the lower performance of tier-to-tier systems, however, we assume that the throughput of the retrieval transactions is  $\lambda_R = 100 \frac{1}{h}$ . The results of the tier-to-tier configuration are presented in appendix C.

### 6.2.1 Design of SBS/RSs under multiple constraints

Given the requirements such as capacity, throughput, utilization and service level, the design decision mainly depends on the annualized costs of the SBS/RS (see also Marchet et al. (2013)), i.e. the system with the lowest costs should be chosen out of the set of system configurations that fulfill the requirements. The costs that have to be considered are the investment costs for the equipment (rack system, lifts, vehicles) and the costs for the system footprint, i.e. the used floor space.

The selection process of the best system configuration can be described as follows. Based on the physical design constraints (maximum length, height and width), we first determine the set of possible system configurations. Afterward, we use the decomposition approach to compute the performance measures such as utilization and retrieval transaction time distribution for all possible configurations. Out of the systems that fulfill the utilization and service level constraints, we choose the one that leads to the lowest costs. Similar approaches have also been used by Marchet et al. (2013) and Epp et al. (2017). In the following, we demonstrate the approach on the given tier-captive example.

In order to determine the set of possible system designs that fulfill the physical constraints, we first compute the maximum number of aisles, tiers and columns that a configuration might have. The maximum number of aisles depends on the width of an aisle, which we assume to be  $d_a = 2m$ , and the maximum width of the system. Hence, we divide the maximum width of the system by the width of an aisle and round down to the next integer value.

$$n_{a,max} = \left\lfloor \frac{max_W}{d_a} \right\rfloor \quad (6.5)$$

In analogy, we compute the maximum number of storage columns.

$$n_{c,max} = \left\lfloor \frac{max_L}{d_c} \right\rfloor \quad (6.6)$$

The maximum number of tiers depends on the maximum height of the system and the number of levels that a vehicle can serve. We assume that all vehicles of the configuration are of the same type, i.e. serving the same amount of levels. Thus, the maximum number of tiers is computed as follows:

$$n_{t,max} = \left\lfloor \frac{max_H}{n_l \cdot d_l} \right\rfloor \quad (6.7)$$

This results in a maximum number of 5 aisles, 27 tiers, and 200 columns in our application example. For a given number of levels per tier  $n_l$ , the minimum number of aisles  $n_{a,min}(n_l)$  needed to reach the required storage capacity is determined by dividing the required number of storage locations  $Cap$  by the maximum number of storage locations that an aisle can hold, which is equal to  $2 \cdot n_{c,max} \cdot n_{t,max} \cdot n_l$ . Since the number of aisles needs to be an integer value, we round the result up to the next integer value.

$$n_{a,min}(n_l) = \left\lceil \frac{Cap}{2 \cdot n_{c,max} \cdot n_{t,max} \cdot n_l} \right\rceil \quad (6.8)$$

In the example with 1 level per tier, we need at least  $n_{a,min} = 2$  aisles to reach the required capacity  $Cap$ . For a given number of aisles  $n_a$  and levels per tier  $n_l$ , the minimum number of tiers  $n_{t,min}(n_a, n_l)$  needed to fulfill the capacity constraint is determined by dividing the required number of storage locations  $Cap$  by the maximum number of storage locations the tiers of the given aisles can hold, which is equal to  $n_a \cdot 2 \cdot n_{c,max} \cdot n_l$ . Again, we round the result up to the next integer value.

$$n_{t,min}(n_a, n_l) = \left\lceil \frac{Cap}{n_a \cdot 2 \cdot n_{c,max} \cdot n_l} \right\rceil \quad (6.9)$$

Given 2 aisles, 1 level per tier and a maximum number of 200 columns, we need at least 25 tiers to reach the required capacity. For any possible combination of aisles and tiers, we determine the number of columns needed on either side of the aisle such that we reach the required capacity, but do not create unnecessary costs in terms of overcapacity. Thus, we divide the required number of storage locations  $Cap$  by the given number of aisles, tiers and levels. Additionally, we divide this value by 2 since we want to determine the number of columns on either side of the aisle. Again, we round the result up to the next integer value.

$$n_c(n_a, n_t, n_l) = \left\lceil \frac{Cap}{n_a \cdot n_t \cdot n_l \cdot 2} \right\rceil \quad (6.10)$$

For example, given 2 aisles and 25 tiers with 1 level per tier, we need 200 columns on either side of the aisle to reach a storage capacity of 20,000 storage locations.

This procedure leads in the application example to the set of system configurations that fulfill the physical requirements. The resulting 47 configurations and their performance measures are presented in table 6.7. The first two columns indicate the configuration number and the system configuration in terms of number of aisles, levels per tier, tiers, and columns on either side of the aisle. The resulting capacity  $Cap = n_a \cdot n_t \cdot n_l \cdot 2 \cdot n_c$  is given in column 3. The footprint of the system (without the lifts, buffers and the conveyor system in front of the aisles) is displayed in column 4. It is determined by the multiplication of the aisle width and aisle length.

$$Footprint = d_a \cdot n_a \cdot d_c \cdot n_c \quad (6.11)$$

Given the number of aisles and tiers, we directly obtain the total number of incoming and outgoing lifts  $n_{L,tot} = 2 \cdot n_a$  (see column 5) and the number of vehicles  $n_V = n_a \cdot n_t$  (see column 6). Finally, the maximum utilization  $\rho_{max,l} = \max\{\rho_{lift_{in}}, \rho_{lift_{out}}\}$  of the lifts and the utilization of the vehicles  $\rho_{veh}$  as well as the 95% quantile of the retrieval transaction time  $T_{R,0.95}$ , which is determined by using the decomposition approach, are presented in columns 7-9, respectively. As in the validation, in case of a utilization larger than 90%, we did not compute the retrieval transaction time distribution. These cases are indicated by n.a.

Table 6.7: Application example single-level shuttle

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [ $m^2$ ]	$n_{L,tot}$	$n_V$	$\rho_{max,l}$	$\rho_{veh}$	$T_{R,0.95}$ [s]
1	2, 1, 25, 200	20,000	400.00	4	50	1.14	0.56	n.a.
2	2, 1, 26, 193	20,072	386.00	4	52	1.15	0.52	n.a.
3	2, 1, 27, 186	20,088	372.00	4	54	1.16	0.49	n.a.
4	3, 1, 17, 197	20,094	591.00	6	51	0.69	0.54	220
5	3, 1, 18, 186	20,088	558.00	6	54	0.70	0.49	190
6	3, 1, 19, 176	20,064	528.00	6	57	0.71	0.44	170
7	3, 1, 20, 167	20,040	501.00	6	60	0.72	0.40	155
8	3, 1, 21, 159	20,034	477.00	6	63	0.73	0.37	143
9	3, 1, 22, 152	20,064	456.00	6	66	0.73	0.34	134
10	3, 1, 23, 145	20,010	435.00	6	69	0.74	0.31	127
11	3, 1, 24, 139	20,016	417.00	6	72	0.75	0.29	122
12	<b>3, 1, 25, 134</b>	<b>20,100</b>	<b>402.00</b>	<b>6</b>	<b>75</b>	<b>0.76</b>	<b>0.27</b>	<b>118</b>
13	<b>3, 1, 26, 129</b>	<b>20,124</b>	<b>387.00</b>	<b>6</b>	<b>78</b>	<b>0.77</b>	<b>0.25</b>	<b>114</b>
14	<b>3, 1, 27, 124</b>	<b>20,088</b>	<b>372.00</b>	<b>6</b>	<b>81</b>	<b>0.77</b>	<b>0.24</b>	<b>111</b>
15	4, 1, 13, 193	20,072	772.00	8	52	0.49	0.52	204
16	4, 1, 14, 179	20,048	716.00	8	56	0.50	0.46	170
17	4, 1, 15, 167	20,040	668.00	8	60	0.51	0.40	148
18	4, 1, 16, 157	20,096	628.00	8	64	0.51	0.36	132
19	4, 1, 17, 148	20,128	592.00	8	68	0.52	0.32	121
20	<b>4, 1, 18, 139</b>	<b>20,016</b>	<b>556.00</b>	<b>8</b>	<b>72</b>	<b>0.53</b>	<b>0.29</b>	<b>111</b>
21	<b>4, 1, 19, 132</b>	<b>20,064</b>	<b>528.00</b>	<b>8</b>	<b>76</b>	<b>0.53</b>	<b>0.27</b>	<b>103</b>
22	<b>4, 1, 20, 125</b>	<b>20,000</b>	<b>500.00</b>	<b>8</b>	<b>80</b>	<b>0.54</b>	<b>0.24</b>	<b>97</b>
23	<b>4, 1, 21, 120</b>	<b>20,160</b>	<b>480.00</b>	<b>8</b>	<b>84</b>	<b>0.54</b>	<b>0.22</b>	<b>92</b>
24	<b>4, 1, 22, 114</b>	<b>20,064</b>	<b>456.00</b>	<b>8</b>	<b>88</b>	<b>0.55</b>	<b>0.21</b>	<b>88</b>
25	<b>4, 1, 23, 109</b>	<b>20,056</b>	<b>436.00</b>	<b>8</b>	<b>92</b>	<b>0.56</b>	<b>0.19</b>	<b>84</b>
26	<b>4, 1, 24, 105</b>	<b>20,160</b>	<b>420.00</b>	<b>8</b>	<b>96</b>	<b>0.56</b>	<b>0.18</b>	<b>82</b>
27	<b>4, 1, 25, 100</b>	<b>20,000</b>	<b>400.00</b>	<b>8</b>	<b>100</b>	<b>0.57</b>	<b>0.16</b>	<b>79</b>
28	<b>4, 1, 26, 97</b>	<b>20,176</b>	<b>388.00</b>	<b>8</b>	<b>104</b>	<b>0.57</b>	<b>0.15</b>	<b>77</b>
29	<b>4, 1, 27, 93</b>	<b>20,088</b>	<b>372.00</b>	<b>8</b>	<b>108</b>	<b>0.58</b>	<b>0.14</b>	<b>75</b>
30	5, 1, 10, 200	20,000	1000.00	10	50	0.38	0.56	225
31	5, 1, 11, 182	20,020	910.00	10	55	0.38	0.47	175
32	5, 1, 12, 167	20,040	835.00	10	60	0.39	0.40	146
33	5, 1, 13, 154	20,020	770.00	10	65	0.39	0.35	127
34	<b>5, 1, 14, 143</b>	<b>20,020</b>	<b>715.00</b>	<b>10</b>	<b>70</b>	<b>0.40</b>	<b>0.31</b>	<b>113</b>
35	<b>5, 1, 15, 134</b>	<b>20,100</b>	<b>670.00</b>	<b>10</b>	<b>75</b>	<b>0.40</b>	<b>0.27</b>	<b>103</b>
36	<b>5, 1, 16, 125</b>	<b>20,000</b>	<b>625.00</b>	<b>10</b>	<b>80</b>	<b>0.41</b>	<b>0.24</b>	<b>94</b>
37	<b>5, 1, 17, 118</b>	<b>20,060</b>	<b>590.00</b>	<b>10</b>	<b>85</b>	<b>0.42</b>	<b>0.22</b>	<b>87</b>
38	<b>5, 1, 18, 112</b>	<b>20,160</b>	<b>560.00</b>	<b>10</b>	<b>90</b>	<b>0.42</b>	<b>0.20</b>	<b>82</b>
39	<b>5, 1, 19, 106</b>	<b>20,140</b>	<b>530.00</b>	<b>10</b>	<b>95</b>	<b>0.43</b>	<b>0.18</b>	<b>78</b>
40	<b>5, 1, 20, 100</b>	<b>20,000</b>	<b>500.00</b>	<b>10</b>	<b>100</b>	<b>0.43</b>	<b>0.16</b>	<b>74</b>

<b>41</b>	<b>5, 1, 21, 96</b>	<b>20,160</b>	<b>480.00</b>	<b>10</b>	<b>105</b>	<b>0.44</b>	<b>0.15</b>	<b>71</b>
<b>42</b>	<b>5, 1, 22, 91</b>	<b>20,020</b>	<b>455.00</b>	<b>10</b>	<b>110</b>	<b>0.44</b>	<b>0.14</b>	<b>68</b>
<b>43</b>	<b>5, 1, 23, 87</b>	<b>20,010</b>	<b>435.00</b>	<b>10</b>	<b>115</b>	<b>0.45</b>	<b>0.13</b>	<b>66</b>
<b>44</b>	<b>5, 1, 24, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>10</b>	<b>120</b>	<b>0.45</b>	<b>0.12</b>	<b>65</b>
<b>45</b>	<b>5, 1, 25, 80</b>	<b>20,000</b>	<b>400.00</b>	<b>10</b>	<b>125</b>	<b>0.45</b>	<b>0.11</b>	<b>63</b>
<b>46</b>	<b>5, 1, 26, 77</b>	<b>20,020</b>	<b>385.00</b>	<b>10</b>	<b>130</b>	<b>0.46</b>	<b>0.11</b>	<b>61</b>
<b>47</b>	<b>5, 1, 27, 75</b>	<b>20,250</b>	<b>375.00</b>	<b>10</b>	<b>135</b>	<b>0.46</b>	<b>0.10</b>	<b>61</b>

We see that the configurations with 2 aisles do not fulfill the requirement of a maximum utilization of less than 90%. Furthermore, only the configurations with 3 aisles and 25-27 tiers, 4 aisles and 18-27 tiers, as well as 5 aisles and 14-27 tiers have a 95% quantile of the retrieval transaction time lower than 120s (printed in bold letters).

In general, the system designer aims at minimizing the costs. In case that the equipment cost has a significant higher impact on the total system cost than the cost for space, the configurations that lead to a low number of lifts and vehicles are of interest, i.e. configurations 12, 20, and 34 given 3, 4, and 5 aisles, respectively. Out of these configurations, configuration 12 has the lowest number of lifts (6), but the highest number of vehicles (75). On the other side, configuration 34 has the highest number of lifts (10), but the lowest number of vehicles (70). Configuration 20 is in between those two configurations with 8 lifts and 72 vehicles.

In contrast, if the cost of space has a significant higher impact on the total system cost than the equipment cost, the system designer would choose configuration 14. In this configuration, the system has the smallest footprint. Additionally, it consists of fewer lifts and vehicles than configurations 29 and 47, which have a similar footprint.

Given the cost structure presented in Marchet et al. (2013), i.e. 50€ per  $m^2$  footprint and year, 10 years of service, 10% interest rate, an investment cost of 10,000€ per vehicle, 50,000€ per lift, and 30€ per storage location, the system designer should choose configuration 12, which leads to a total annualized system cost of 289,118€ (incl. the discounted investment costs).

The results of the tier-to-tier example are presented in appendix C in table C.2. The configurations with 2-4 aisles do not meet the system requirement in terms of the 95% quantile. Only in the configurations with 5 aisles and 20-27 tiers, the 95% quantile of the retrieval transaction time is less than 120s (printed in

bold letters). All of these configurations have 5 lifts and 5 vehicles. Thus, the system designer may choose the configuration with the smallest footprint, i.e. the configuration with the most tiers (configuration 47).

The results of the tier-captive configurations also indicate that the utilization of the vehicles is lower than the maximum utilization of the lifts in most cases. In general, it is lower than 60% in all analyzed configurations. This leads to the question whether it is beneficial to reduce the number of needed vehicles by increasing their utilization. The utilization of the vehicles can be increased by using multi-level shuttles. Hence, we investigate the impact of vehicles that can serve more than one level on the performance measures in the following section.

### 6.2.2 Impact of multi-level shuttles

Given the physical constraints, we determine the possible configurations with multi-level shuttles in analogy to the configurations with single-level shuttles. The results of the tier-captive configuration for two, three, and four levels per vehicle are presented in tables 6.8, 6.9, and 6.10, respectively.

We see that the multi-level shuttles have a large impact on the design decision. In configurations with the same number of aisle and vertical/horizontal storage locations (e.g., configuration 13 and configuration 53), the systems with multi-level shuttles lead to a higher utilization of the vehicles, thus a longer retrieval transaction time. Whereas configuration 13 fulfills all requirements, configuration 53 has a too large 95% quantile of the retrieval transaction time. On the other side, multi-level shuttles can lead in configurations, in which the utilization of the single-level shuttles is low, to an increase in vehicle utilization without violating the requirements. For example, configuration 44 with 10 lifts and 120 single-level shuttles (vehicle utilization of 12%) leads to a 95% quantile of the retrieval transaction time of 65s, whereas configuration 94 with 10 lifts and only 30 multi-level shuttles that can serve 4 levels (vehicle utilization of 49%) leads to a 95% quantile of the retrieval transaction time of 100s, thus still under 120s.

Table 6.8: Application example multi-level shuttle with two levels

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [ $m^2$ ]	$n_{L,tot}$	$n_V$	$\rho_{max,l}$	$\rho_{veh}$	$T_{R,0.95}$ [s]
48	2, 2, 13, 193	20,072	386.00	4	26	1.13	1.05	n.a.
49	3, 2, 9, 186	20,088	558.00	6	27	0.69	0.98	n.a.
50	3, 2, 10, 167	20,040	501.00	6	30	0.71	0.81	423
51	3, 2, 11, 152	20,064	456.00	6	33	0.72	0.68	246
52	3, 2, 12, 139	20,016	417.00	6	36	0.74	0.58	183
53	3, 2, 13, 129	20,124	387.00	6	39	0.76	0.51	153
54	4, 2, 7, 179	20,048	716.00	8	28	0.49	0.91	n.a.
55	4, 2, 8, 157	20,096	628.00	8	32	0.51	0.72	280
56	4, 2, 9, 139	20,016	556.00	8	36	0.52	0.58	176
57	4, 2, 10, 125	20,000	500.00	8	40	0.53	0.48	135
<b>58</b>	<b>4, 2, 11, 114</b>	<b>20,064</b>	<b>456.00</b>	<b>8</b>	<b>44</b>	<b>0.54</b>	<b>0.41</b>	<b>113</b>
<b>59</b>	<b>4, 2, 12, 105</b>	<b>20,160</b>	<b>420.00</b>	<b>8</b>	<b>48</b>	<b>0.56</b>	<b>0.35</b>	<b>99</b>
<b>60</b>	<b>4, 2, 13, 97</b>	<b>20,176</b>	<b>388.00</b>	<b>8</b>	<b>52</b>	<b>0.57</b>	<b>0.31</b>	<b>90</b>
61	5, 2, 5, 200	20,000	1000.00	10	25	0.37	1.12	n.a.
62	5, 2, 6, 167	20,040	835.00	10	30	0.38	0.81	419
63	5, 2, 7, 143	20,020	715.00	10	35	0.39	0.61	191
64	5, 2, 8, 125	20,000	625.00	10	40	0.40	0.48	133
<b>65</b>	<b>5, 2, 9, 112</b>	<b>20,160</b>	<b>560.00</b>	<b>10</b>	<b>45</b>	<b>0.42</b>	<b>0.40</b>	<b>107</b>
<b>66</b>	<b>5, 2, 10, 100</b>	<b>20,000</b>	<b>500.00</b>	<b>10</b>	<b>50</b>	<b>0.43</b>	<b>0.33</b>	<b>90</b>
<b>67</b>	<b>5, 2, 11, 91</b>	<b>20,020</b>	<b>455.00</b>	<b>10</b>	<b>55</b>	<b>0.43</b>	<b>0.28</b>	<b>80</b>
<b>68</b>	<b>5, 2, 12, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>10</b>	<b>60</b>	<b>0.44</b>	<b>0.24</b>	<b>73</b>
<b>69</b>	<b>5, 2, 13, 77</b>	<b>20,020</b>	<b>385.00</b>	<b>10</b>	<b>65</b>	<b>0.45</b>	<b>0.21</b>	<b>68</b>

In the given example, all configurations with multi-level shuttles that fulfill the requirements (configurations 58-60, 65-69, 79, 83-85, and 94) need a smaller number of vehicles than the configurations with single-level shuttles that fulfill the constraints. For example, configuration 94 only needs 30 shuttles to fulfill the given requirements, whereas the smallest number of shuttles needed with single-level shuttles is 70 (configuration 34).

Hence, the system designer may choose the configuration that leads to the lowest costs. Given a small difference in costs for single-level and multi-level shuttles, it is obvious that multi-level shuttles offer an interesting option when deciding about the type of system during the design phase of an SBS/RS.

Table 6.9: Application example multi-level shuttle with three levels

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [ $m^2$ ]	$n_{L,tot}$	$n_V$	$\rho_{max,l}$	$\rho_{veh}$	$T_{R,0.95}$ [s]
70	2, 3, 9, 186	20,088	372.00	4	18	1.14	1.47	n.a.
71	3, 3, 6, 186	20,088	558.00	6	18	0.69	1.47	n.a.
72	3, 3, 7, 159	20,034	477.00	6	21	0.72	1.11	n.a.
73	3, 3, 8, 139	20,016	417.00	6	24	0.74	0.87	541
74	3, 3, 9, 124	20,088	372.00	6	27	0.76	0.71	232
75	4, 3, 5, 167	20,040	668.00	8	20	0.50	1.21	n.a.
76	4, 3, 6, 139	20,016	556.00	8	24	0.52	0.87	540
77	4, 3, 7, 120	20,160	480.00	8	28	0.54	0.67	194
78	4, 3, 8, 105	20,160	420.00	8	32	0.55	0.53	129
<b>79</b>	<b>4, 3, 9, 93</b>	<b>20,088</b>	<b>372.00</b>	<b>8</b>	<b>36</b>	<b>0.57</b>	<b>0.43</b>	<b>102</b>
80	5, 3, 4, 167	20,040	835.00	10	20	0.39	1.21	n.a.
81	5, 3, 5, 134	20,100	670.00	10	25	0.40	0.82	363
82	5, 3, 6, 112	20,160	560.00	10	30	0.41	0.59	152
<b>83</b>	<b>5, 3, 7, 96</b>	<b>20,160</b>	<b>480.00</b>	<b>10</b>	<b>35</b>	<b>0.43</b>	<b>0.46</b>	<b>105</b>
<b>84</b>	<b>5, 3, 8, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>10</b>	<b>40</b>	<b>0.44</b>	<b>0.36</b>	<b>84</b>
85	5, 3, 9, 75	20,250	375.00	10	45	0.46	0.30	73

Table 6.10: Application example multi-level shuttle with four levels

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [ $m^2$ ]	$n_{L,tot}$	$n_V$	$\rho_{max,l}$	$\rho_{veh}$	$T_{R,0.95}$ [s]
86	3, 4, 5, 167	20,040	501.00	6	15	0.70	1.61	n.a.
87	3, 4, 6, 139	20,016	417.00	6	18	0.74	1.17	n.a.
88	4, 4, 4, 157	20,096	628.00	8	16	0.50	1.44	n.a.
89	4, 4, 5, 125	20,000	500.00	8	20	0.53	0.97	n.a.
90	4, 4, 6, 105	20,160	420.00	8	24	0.55	0.71	196
91	5, 4, 3, 167	20,040	835.00	10	15	0.39	1.61	n.a.
92	5, 4, 4, 125	20,000	625.00	10	20	0.40	0.97	n.a.
93	5, 4, 5, 100	20,000	500.00	10	25	0.42	0.66	162
<b>94</b>	<b>5, 4, 6, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>10</b>	<b>30</b>	<b>0.44</b>	<b>0.49</b>	<b>100</b>

In the case of the tier-to-tier configurations with only one vehicle per aisle (presented in appendix C in tables C.3, C.4, and C.5), the usage of multi-level shuttles does result for the same physical configurations (i.e. the same number

of aisles and storage locations in vertical and horizontal direction) in a lower utilization of the aisle. Since the vehicle changes the tiers less often, the service time of the aisle station decreases. However, the reduction of the utilization is too low to lead to a different preferable system design of the application example. We see that we need 5 aisles, i.e. 5 lifts and 5 vehicles, to fulfill the constraint that the 95% quantile of the retrieval transaction time is at maximum 120s, no matter how many levels a vehicle can serve. We see that the configurations which fulfill the constraints need to have a rather low utilization to meet the 95% quantile of the retrieval transaction time. We know that different utilization levels have a rather low impact on the sojourn time of a queueing system given low utilization levels (in comparison to high utilization levels). Hence, the chosen requirements only lead to small differences in the performance measures of the given multi-level shuttles. As a result, the decision in our example only depends on the cost for space and multi-level shuttles. The configurations that fulfill all requirements with the smallest footprint are configuration 47 (single-level shuttle,  $T_{R,0.95} = 87s$ ) and configuration 85 (multi-level shuttle with 3 levels,  $T_{R,0.95} = 85s$ ). If we assume that the single-level shuttles are cheaper than the multi-level shuttles, the system designer should choose configuration 47.

### 6.2.3 Impact of re-entrant bins and variability of random variables

In this section, we show the impact of re-entrant bins and different levels of variability on the design decision. To do this, we use the system design example of the previous section, i.e. we assume the same resource characteristics, control policies, and design requirements. Additionally, we consider multi-level shuttles. This results in the same 94 configurations that fulfill the physical constraints as in the previous sections.

In order to show the impact of re-entrant bins and different levels of variability on the design decision, we compute the performance measures of the 94 configurations varying the following parameters.

Table 6.11: Varied parameters of the tier-captive and tier-to-tier SBS/RSS

Parameter	Values
$p_{ps}$	{0.0, 1.0}
$p_{em}$	0.0
$B_{ps}$	{Exponential, Gamma with $cv^2 = 0.025$ }
$A_{S_{rep}}$	{Exponential, Gamma with $cv^2 = 0.025$ }
$A_R$	{Exponential, Gamma with $cv^2 = 0.025$ }

This means that we investigate the impact of re-entrant bins on the design decision by setting the probability of an empty bin after picking to  $p_{em} = 0.0$  and varying the picking probability  $p_{ps}$  from 0.0 to 1.0. The probability  $p_{ps} = 0.0$  corresponds to the case that there is no re-entrance since all bins bypass the picking stations and leave the system after being retrieved. In contrast, the probability  $p_{ps} = 1.0$  results in a re-entrance of all retrieved bins due to  $p_{em} = 0.0$ .

If there is no re-entrance, the random variables of interest are the retrieval inter-arrival time  $A_R$  and the inter-arrival time of replenishment storage transactions  $A_{S_{rep}}$ . Both may be distributed according to an Exponential distribution with an scv of 1.0, which corresponds to random arrival processes, or to a Gamma distribution with an scv of 0.025, which corresponds to stable arrival processes. The picking process is of no interest since there is no picking in the case of no re-entrance.

If all bins re-enter the system after picking, the random variables of interest are the retrieval inter-arrival time  $A_R$  and the service time of the picking stations  $B_{ps}$ . Again, both random variables may be distributed according to an Exponential distribution ( $cv^2 = 1.0$ ) or to a Gamma distribution ( $cv^2 = 0.025$ ). Since there is no replenishment of bins in this case, the replenishment arrival process is of no interest.

In total, this results in a full factorial experiment that consists of the following 8 combinations. If a random variable does not exist in a given combination, we indicate this by n.a.

Table 6.12: Possible combinations of the tier-captive and tier-to-tier SBS/RSSs

Combi.	re-entrance ( $p_{ps}$ )	$cv^2[A_R]$	$cv^2[A_{S_{rep}}]$	$cv^2[B_{ps}]$
1	no (0.0)	1.0	1.0	n.a.
2	no (0.0)	1.0	0.025	n.a.
3	no (0.0)	0.025	1.0	n.a.
4	no (0.0)	0.025	0.025	n.a.
5	yes (1.0)	1.0	n.a.	1.0
6	yes (1.0)	1.0	n.a.	0.025
7	yes (1.0)	0.025	n.a.	1.0
8	yes (1.0)	0.025	n.a.	0.025

For each combination, we compute the 95% quantile of the retrieval transaction time of the 94 configurations that fulfill the physical constraints. The impact of the variability reduction of the retrieval inter-arrival time distribution on the 95% quantile of the retrieval transaction time is determined by comparing the results of combination 1 with combination 3, 2 with 4, 5 with 7, and 6 with 8, respectively. Hereby, we evaluate the differences in the performance measures for the same configuration by only changing the scv of  $A_R$  from 1.0 to 0.025 (the other parameters and random variables remain unchanged). For every configuration  $c$  that has a utilization below 90%, we compute the average relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the retrieval inter-arrival time variability as follows.

$$\bar{\Delta}_{cv^2[A_R],c}^{rel} = \frac{1}{4} \cdot \left( \frac{T_{R,0.95,c}^{combi.1} - T_{R,0.95,c}^{combi.3}}{T_{R,0.95,c}^{combi.1}} + \frac{T_{R,0.95,c}^{combi.2} - T_{R,0.95,c}^{combi.4}}{T_{R,0.95,c}^{combi.2}} \right. \\ \left. + \frac{T_{R,0.95,c}^{combi.5} - T_{R,0.95,c}^{combi.7}}{T_{R,0.95,c}^{combi.5}} + \frac{T_{R,0.95,c}^{combi.6} - T_{R,0.95,c}^{combi.8}}{T_{R,0.95,c}^{combi.6}} \right) \quad (6.12)$$

The maximum and minimum reductions caused by the reduction of the retrieval arrival stream variability ( $\Delta_{cv^2[A_R],c}^{rel,max}$  and  $\Delta_{cv^2[A_R],c}^{rel,min}$ ) are obtained by the maximum and minimum value of the summands within the bracket of equation 6.12.

Moreover, we compute the overall average, maximum and minimum relative reduction of  $T_{R,0.95}$  given the results of all  $n_{config}^{nev}$  configurations that have a utilization below 90%.

$$\bar{\Delta}_{cv^2[A_R],all}^{rel} = \frac{1}{n_{config}^{nev}} \cdot \sum_{c=1}^{n_{config}^{nev}} \bar{\Delta}_{cv^2[A_R],c}^{rel} \quad (6.13)$$

$$\Delta_{cv^2[A_R],all}^{rel,max} = \max \{ \Delta_{cv^2[A_R],1}^{rel,max}; \dots; \Delta_{cv^2[A_R],n_{config}^{nev}}^{rel,max} \} \quad (6.14)$$

$$\Delta_{cv^2[A_R],all}^{rel,min} = \min \{ \Delta_{cv^2[A_R],1}^{rel,min}; \dots; \Delta_{cv^2[A_R],n_{config}^{nev}}^{rel,min} \} \quad (6.15)$$

In analogy, we compute the impact of the variability reduction of the arrival process of the storage transactions for replenishment (equations 6.16 - 6.19) and the service process of the picking stations (equations 6.20 - 6.23) on the 95% quantile of the retrieval transaction time.

$$\bar{\Delta}_{cv^2[A_{Srep}],c}^{rel} = \frac{1}{2} \cdot \left( \frac{T_{R,0.95,c}^{combi,1} - T_{R,0.95,c}^{combi,2}}{T_{R,0.95,c}^{combi,1}} + \frac{T_{R,0.95,c}^{combi,3} - T_{R,0.95,c}^{combi,4}}{T_{R,0.95,c}^{combi,3}} \right) \quad (6.16)$$

$$\bar{\Delta}_{cv^2[A_{Srep}],all}^{rel} = \frac{1}{n_{config}^{nev}} \cdot \sum_{c=1}^{n_{config}^{nev}} \bar{\Delta}_{cv^2[A_{Srep}],c}^{rel} \quad (6.17)$$

$$\Delta_{cv^2[A_{Srep}],all}^{rel,max} = \max \{ \Delta_{cv^2[A_{Srep}],1}^{rel,max}; \dots; \Delta_{cv^2[A_{Srep}],n_{config}^{nev}}^{rel,max} \} \quad (6.18)$$

$$\Delta_{cv^2[A_{Srep}],all}^{rel,min} = \min \{ \Delta_{cv^2[A_{Srep}],1}^{rel,min}; \dots; \Delta_{cv^2[A_{Srep}],n_{config}^{nev}}^{rel,min} \} \quad (6.19)$$

$$\bar{\Delta}_{cv^2[B_{ps}],c}^{rel} = \frac{1}{2} \cdot \left( \frac{T_{R,0.95,c}^{combi,5} - T_{R,0.95,c}^{combi,6}}{T_{R,0.95,c}^{combi,5}} + \frac{T_{R,0.95,c}^{combi,7} - T_{R,0.95,c}^{combi,8}}{T_{R,0.95,c}^{combi,7}} \right) \quad (6.20)$$

$$\bar{\Delta}_{cv^2[B_{ps}],all}^{rel} = \frac{1}{n_{config}^{nev}} \cdot \sum_{c=1}^{n_{config}^{nev}} \bar{\Delta}_{cv^2[B_{ps}],c}^{rel} \quad (6.21)$$

$$\Delta_{cv^2[B_{ps}],all}^{rel,max} = \max \{ \Delta_{cv^2[B_{ps}],1}^{rel,max}; \dots; \Delta_{cv^2[B_{ps}],n_{config}^{nev}}^{rel,max} \} \quad (6.22)$$

$$\Delta_{cv^2[B_{ps}],all}^{rel,min} = \min \{ \Delta_{cv^2[B_{ps}],1}^{rel,min}; \dots; \Delta_{cv^2[B_{ps}],n_{config}^{nev}}^{rel,min} \} \quad (6.23)$$

To evaluate the impact of re-entrant bins, we compare the results of combinations 5 and 6 with the results of combinations 1 and 2 ( $cv^2[A_R] = 1.0$ ) as well as the results of combinations 7 and 8 with the results of combinations 3 and 4 ( $cv^2[A_R] = 0.025$ ).

$$\begin{aligned} \bar{\Delta}_{noRe-en,c}^{rel} = & \frac{1}{8} \cdot \left( \frac{T_{R,0.95,c}^{combi.5} - T_{R,0.95,c}^{combi.1}}{T_{R,0.95,c}^{combi.5}} + \frac{T_{R,0.95,c}^{combi.6} - T_{R,0.95,c}^{combi.1}}{T_{R,0.95,c}^{combi.6}} \right. \\ & + \frac{T_{R,0.95,c}^{combi.5} - T_{R,0.95,c}^{combi.2}}{T_{R,0.95,c}^{combi.5}} + \frac{T_{R,0.95,c}^{combi.6} - T_{R,0.95,c}^{combi.2}}{T_{R,0.95,c}^{combi.6}} \\ & + \frac{T_{R,0.95,c}^{combi.7} - T_{R,0.95,c}^{combi.3}}{T_{R,0.95,c}^{combi.7}} + \frac{T_{R,0.95,c}^{combi.8} - T_{R,0.95,c}^{combi.3}}{T_{R,0.95,c}^{combi.8}} \\ & \left. + \frac{T_{R,0.95,c}^{combi.7} - T_{R,0.95,c}^{combi.4}}{T_{R,0.95,c}^{combi.7}} + \frac{T_{R,0.95,c}^{combi.8} - T_{R,0.95,c}^{combi.4}}{T_{R,0.95,c}^{combi.8}} \right) \end{aligned} \quad (6.24)$$

$$\bar{\Delta}_{noRe-en,all}^{rel} = \frac{1}{n_{config}^{nev}} \cdot \sum_{c=1}^{n_{config}^{nev}} \bar{\Delta}_{noRe-en,c}^{rel} \quad (6.25)$$

$$\Delta_{noRe-en,all}^{rel,max} = \max \{ \Delta_{noRe-en,1}^{rel,max}; \dots; \Delta_{noRe-en,n_{config}^{nev}}^{rel,max} \} \quad (6.26)$$

$$\Delta_{noRe-en,all}^{rel,min} = \min \{ \Delta_{noRe-en,1}^{rel,min}; \dots; \Delta_{noRe-en,n_{config}^{nev}}^{rel,min} \} \quad (6.27)$$

The resulting 95% quantiles of the retrieval transaction time of all 8 combinations and 94 tier-captive configurations are displayed in appendix B in table B.1. Moreover, the results of the 94 tier-to-tier configurations are displayed in appendix C in table C.6. Configurations with a utilization above 90% are

indicated by n.a. The aggregated results of the tier-captive and tier-to-tier configurations are presented in table 6.13 and the following sections.

Table 6.13: Impact of the variability reduction of random variables and the absence of re-entrant bins on the 95% quantile of the retrieval transaction time

Parameter	tier-captive			tier-to-tier		
	min	avg ( $\bar{\Delta}$ )	max	min	avg ( $\bar{\Delta}$ )	max
$\Delta_{cv^2[A_R],all}^{rel}$	0.00%	0.59%	1.64%	6.45%	8.67%	12.90%
$\Delta_{cv^2[A_{Srep}],all}^{rel}$	0.00%	0.21%	1.47%	6.67%	8.63%	12.90%
$\Delta_{cv^2[B_{ps}],all}^{rel}$	0.00%	0.04%	0.68%	0.47%	1.14%	2.44%
$\Delta_{noRe-en,all}^{rel}$	0.00%	0.14%	1.47%	-3.94%	2.84%	10.87%

### Impact on the tier-captive configurations

In the case of the tier-captive systems, the different combinations lead approximately to the same 95% quantiles of the retrieval transaction time for a given configuration. Reasoning for this can be found in the behavior of stochastic streams during their split. Due to the random storage assignment rule and the large number of vehicles, the transaction streams are split into a large number of single streams with a small routing probability. Since the split of stochastic streams into many directions with a small routing probability leads to a similar arrival variability ( $cv^2$  tends towards 1.0) at the subsequent queueing systems (see also Whitt (1983) and Furmans (2000)), the inter-arrival time variability at the vehicle stations are similar for all combinations. Since the service time distributions are identical for all combinations, this leads to a similar queueing behavior at the vehicle station and outgoing lift station. Hence, the design decision given the design requirements does not depend on the proportion of re-entrant bins or the variability of the random variables in the given example. An exception are the cases, in which the 95% quantile of the retrieval transaction time assumes either 120s or 121s. Here, the decision on the system design may change.

In the application example, the overall average relative reduction of  $T_{R,0.95}$  caused by the reduction of the variability of the random variables  $A_R$ ,  $A_{S_{rep}}$ , and  $B_{ps}$  are equal to  $\bar{\Delta}_{cv^2[A_R],all}^{rel} = 0.59\%$ ,  $\bar{\Delta}_{cv^2[A_{S_{rep}}],all}^{rel} = 0.21\%$ , and  $\bar{\Delta}_{cv^2[B_{ps}],all}^{rel} = 0.04\%$ , respectively. As stated above, the results indicate that the reduction in the retrieval transaction time caused by a reduction of the variability of the random variables is very low due to the large network with many splits of the transaction streams. As expected, in none of the comparisons there was an increase in the retrieval transaction time, i.e. the reduction of the variability of the random variables had a positive effect on the retrieval transaction time.

While evaluating the approximation quality in section 6.1.2, we have seen that the approximation quality of systems with Poisson arrivals is very high regarding the 95% quantile of the retrieval transaction time (on average 0.81% using the stochastic split method), whereas non-Poisson arrivals lead to slightly larger deviations (on average 3.00% using the stochastic split method, see also table 6.2). Moreover, the decomposition approach mostly overestimates the retrieval transaction time (see also figure 6.1). This indicates that the average relative reduction of  $T_{R,0.95}$  caused by the reduction of the variability of the random variables is likely to be underestimated. However, due to the small differences in the average deviations of the 95% quantile of the retrieval transaction times given Poisson and non-Poisson arrivals, the underestimation should be rather low. Hence, the overall impact of the reduction of the arrival stream variability should be rather small in the case of tier-captive systems with many splitting and merging operations.

When comparing the combinations with re-entrant bins to the combinations without re-entrant bins, we see that combinations without re-entrant bins on average lead to a marginal reduction of the 95% quantile of the retrieval transaction time ( $\bar{\Delta}_{noRe-en,all}^{rel} = 0.14\%$ ). Again, there is almost no impact on the performance measures.

Since the re-entrance of bins leads to an increase in computation times, we may ignore the picking process, thus the re-entrance of bins, in large systems with many splitting operations. In the application example, the average computation time using the stochastic split method is 25.00s, if there are no re-entrant bins, and 114.99s, if there are 100% re-entrant bins. In the validation presented in section 6.1.2, the average computation times of a tier-captive configuration using the stochastic split method and the fast split approximation are 36.10s

and 9.27s, respectively. If we only consider the tier-captive configurations with  $p_{ps} = 0.0$ , the average computation times using the stochastic split method and the fast split approximation are 9.83s and 1.30s, respectively.

### Impact on the tier-to-tier configurations

In the case of the tier-to-tier systems, the different combinations lead to a higher difference of the 95% quantiles of the retrieval transaction time for a given configuration. Since the stochastic streams are not split into many streams, the variability of the random variables has a larger impact on the performance measures, i.e. the design decision. In general, the reduction of the variability of the random variables leads to a shorter retrieval transaction time. Hence, system configurations with a higher utilization but a shorter retrieval transaction time due to the reduced variabilities are recommended. For example, in the case of single-level shuttles, the reduction of the retrieval inter-arrival time variability leads to a reduction of the 95% quantile of the retrieval transaction time which makes it possible to operate the system with 4 aisles (vehicles) instead of 5 aisles (vehicles).

In general, the overall average relative reductions of  $T_{R,0.95}$  caused by the reduction of the variability of the random variables  $A_R$ ,  $A_{Srep}$ , and  $B_{ps}$  are equal to  $\bar{\Delta}_{cv^2[A_R],all}^{rel} = 8.67\%$ ,  $\bar{\Delta}_{cv^2[A_{Srep}],all}^{rel} = 8.63\%$ , and  $\bar{\Delta}_{cv^2[B_{ps}],all}^{rel} = 1.14\%$ , respectively. As expected, the reduction of the variabilities does not lead to an increase in the retrieval transaction time. Moreover, the results indicate that the reduction of the inter-arrival time variabilities has a larger impact than the reduction of the service time variability of the picking stations. Since there are many splitting operations in between the picking stations and the re-entrance of the bins, the reduction of the service time variability of the picking stations does have a lower impact than the reduction of the retrieval and storage replenishment inter-arrival time variabilities, which do not undergo any additional process before being split among the aisles.

Furthermore, the reduction of  $T_{R,0.95}$  caused by the reduction of the variability of the random variables should be even higher considering the results of the validation. In section 6.1.2, the results indicate that a small number of aisles and non-Poisson arrivals of the retrieval transactions lead to a large overestimation of the retrieval transaction time (on average 18.51%, see also table

6.4 and figure 6.6). This means that we most likely underestimate the already large impact of the reduction of the variability of the retrieval transaction arrival process. Hence, the reduction of the arrival process variability has a large impact on the performance measures of tier-to-tier systems with a single vehicle per aisle and a small number of aisles, i.e. a small number of splitting and merging operations.

When comparing the combinations with 100% re-entrant bins (i.e. no  $A_{S_{rep}}$ ) to the combinations without re-entrant bins (i.e.  $A_{S_{rep}}$  with  $cv^2[A_{S_{rep}}] = 0.025$  or  $cv^2[A_{S_{rep}}] = 1.0$ ), we see that combinations without re-entrant bins on average lead to a reduction of the 95% quantile of the retrieval transaction time ( $\bar{\Delta}_{noRe-en,all}^{rel} = 2.84\%$ ). However, there are cases in which combinations without re-entrant bins lead to a reduction of  $T_{R,0.95}$ , and there are cases in which they lead to an increase of  $T_{R,0.95}$ . In particular, in the cases with  $cv^2[A_{S_{rep}}] = 0.025$ , the combinations without re-entrant bins lead to an average reduction of  $T_{R,0.95}$  of 7.23%. Furthermore, in these cases, the reduction is not less than 4.94%. On the contrary, in the cases with  $cv^2[A_{S_{rep}}] = 1.0$ , the combinations without re-entrant bins lead to an average increase of  $T_{R,0.95}$  of 1.55%. Moreover, there are no cases with  $cv^2[A_{S_{rep}}] = 1.0$  that lead to a reduction of  $T_{R,0.95}$ . This shows that the behavior of  $T_{R,0.95}$  is mainly driven by the variabilities of the random variables, and not by the existence of re-entrant bins.



# 7 Conclusion

## 7.1 Summary

A shuttle-based storage and retrieval system (SBS/RS) is a special type of automated storage and retrieval system for mini-loads. SBS/RSs can be classified according to the degree of freedom of the vehicles that are installed to store and retrieve the goods. Tier-to-tier vehicles are able to change the tiers by using a lift, whereas tier-captive vehicles stay in their dedicated tier. In analogy, aisle-to-aisle vehicles are able to change the aisles, and aisle-captive vehicles stay in the aisle. Hence, different system configurations can be installed. Whereas tier-captive configurations with vehicles that stay in their aisle can achieve a high throughput, tier-to-tier configurations with vehicles that can change the aisles offer a high flexibility since additional vehicles can be added to match the needed throughput. Thus, shuttle-based storage and retrieval systems are installed more and more frequently in warehouses, especially in part-to-picker systems. This leads to the need for performance evaluation tools that can be used during the design phase of SBS/RSs. In particular, models have to be developed that consider the shape of the processing time distributions and allow the computation of the complete probability distributions of the performance measures.

Hence, the focus of this work was the development of a new approach that can be used to model and analyze different configurations of SBS/RSs. The approach is based on the modeling of the SBS/RSs as discrete-time open queueing networks. Afterward, we decompose the queueing network into independent  $G|G|1$  queueing systems that represent the parts of the system with a population constraint of one due to the use of one resource (e.g., shuttle, lift). We model the service time distributions of the queueing systems on the basis of the type of transactions that are using the resource, the physical design of the system, the characteristics of the resources, and the applied control

policies. Given the inter-arrival time distributions of the incoming storage and retrieval transactions, the routing of the transactions through the network, and the service time distributions of the resources, we compute the distributions of the retrieval transaction time and the number of transactions that are physically waiting in front of the resources. We achieve this by applying the discrete-time methods to compute the sojourn time and inter-departure distributions of a  $G|G|1$  queueing system and the inter-arrival time distributions after the split or merge of stochastic streams. Since the computation of the stochastic streams after a splitting operation can become time consuming, we develop a new approximation method that results in faster computation times, the so-called fast split approximation. Due to the possible re-entrance of transactions after the picking process, an iterative procedure is used to compute the overall performance measures.

Subsequently, the approach is applied to both a tier-captive and a tier-to-tier SBS/RS, in which the vehicles are not able to change the aisles. In the modeled tier-captive system, two lifts per aisle are installed. One lift is used to transport the incoming storage transactions from the input point of the aisle to the target tiers, and the other lift is used to transport the outgoing retrieval transactions from the tiers to the output point of the aisle. In the modeled tier-to-tier system, one vehicle and one lift per aisle are installed. In both systems, the number of aisles, tiers per aisle, and storage columns on either side of the aisle are input parameters. Moreover, the load handling device of the vehicles can reach one or more than one level of the single-deep storage rack. Regarding the system load, the access frequency of the retrieval transactions is assumed to be equal among the storage locations. Regarding the control policies, we assumed a random storage assignment rule, single command cycles, FCFS sequencing, and a POSC dwell point strategy. Both type of systems are connected to a conveyor system that transports the bins to the succeeding picking stations. From there, non-empty bins are routed back to the SBS/RS, whereas empty bins leave the system for replenishment. By applying the decomposition approach to these systems, we are the first that allow the computation of the retrieval transaction time distribution for SBS/RSs with multi-level shuttles, picking stations and re-entrant bins.

Since the decomposition approach is not exact, we analyzed its approximation quality by comparing the performance measures of the decomposition approach to the performance measures that are obtained by a discrete-event simu-

lation. Since the complete probability distributions can be computed, we compared the results based on the expected value and the 95% quantile of the following performance measures: retrieval transaction time, inter-departure time of the transactions leaving the system, storage transaction queue length in front of the lift, and queue length in front of the picking stations. In order to test a large set of different system configurations, we varied for both the tier-captive and tier-to-tier SBS/RS the following system parameters: number of aisles, number of tiers, number of levels a shuttle can reach, number of columns on either side of the aisle, probability that a bin is routed to one of the picking stations, as well as expected value and variability of the inter-arrival time of the incoming retrieval transactions. The data regarding the size of the equally sized storage columns as well as the velocities, acceleration/deceleration rates and transfer times of the lifts and the vehicles were given to us by a European material handling provider. The combination of the input parameters resulted for both the tier-captive and the tier-to-tier SBS/RSs in over 1,000 analyzed system configurations with a maximum utilization of the resources between 50% and 90%.

In general, the decomposition approach reaches a high approximation quality. The average deviations of the performance measures of the retrieval transaction time distribution and the inter-departure time distribution are smaller than 4% for both the tier-captive and the tier-to-tier configurations. Especially in system configurations with Poisson arrivals and a large number of aisles and tiers, the average deviations are very low. For example, in tier-captive configurations with Poisson arrivals, the average deviations are below 1.5%. The same is valid for tier-to-tier configurations with Poisson arrivals and 8 aisles. Regarding the queue length distributions, the average deviations of the performance measures are less than 1 bin. The use of the newly developed fast split approximation results in an average reduction of the computation times by more than 50%. At the same time, the approximation quality using the fast split approximation is just slightly lower than the approximation quality using the stochastic split method.

After the validation, we used the decomposition approach to show how to configure a suitable system design given the following design requirements: storage capacity, maximum floor space and height, throughput, and service level in terms of the 95% quantile of the retrieval transaction time distribution. Moreover, we demonstrated the impact of multi-level shuttles on the design decision.

As expected, less shuttles in the system lead to a higher utilization of the individual shuttles. Since the single-level shuttles were very low utilized in many configurations, the use of multi-level shuttles resulted in system configurations that consisted of less shuttles, but still fulfilled the system requirements. Hence, multi-level shuttles offer an interesting alternative to single-level shuttles if the difference in costs between the single-level and multi-level shuttles is small. In addition, we analyzed the impact of re-entrant bins and the variability of the random variables retrieval transaction inter-arrival time, storage transaction for replenishment inter-arrival time, and picking station service time on the 95% quantile of the retrieval transaction time. In general, the impact of the re-entrant bins and the service time variability of the picking process is rather low. The positive impact of the reduction of the variability of the storage and retrieval arrival processes on the 95% quantile of the retrieval transaction time depends on the complexity of the transaction flow. Whereas the impact is rather low in the case of the analyzed tier-captive systems that contain many splitting and merging operations, the reduction of the 95% quantile of the retrieval transaction time is higher in the case of the analyzed tier-to-tier systems that contain less splitting and merging operations.

## 7.2 Outlook

The presented decomposition approach allows for the computation of the distribution of the performance measures of SBS/RSSs with re-entrant bins, picking stations, and multi-level shuttles. The approximation quality of the approach is high, especially given Poisson arrivals and system configurations with many splitting and merging operations. As a result, the approach can be used during the early planning phase of SBS/RSSs to determine suitable system designs that fulfill the customer requirements. There are several possible extensions regarding the physical design and the applied control policies. For example, we could model SBS/RSSs with double-deep storage locations or a class-based storage assignment rule by adjusting the modeling of the service time distributions. Moreover, by including other discrete-time building blocks such as a  $G|G|m$  queueing system into the modeling approach, we could approximate the behavior of system parts with a population constraint larger than one. This would allow us to analyze systems with more than one tier-to-tier vehi-

cle per aisle or more than one aisle-to-aisle vehicle per tier. A special focus would have to be placed on the appropriate modeling of blocking behavior and deadlock avoiding algorithms. A more appropriate way to model system parts with a population constraint larger than one would be the development of discrete-time methods of semi-open queueing networks. However, the use of these more complex discrete-time methods would lead to an increase of the computation times. Thus, a prerequisite is the development of methods that speed up the computation times of discrete-time queueing systems and networks that are based on discrete-time Markov chains. Additionally, in our approach we model the buffer places in front of the resources as queues with unlimited capacity. To model limited buffer capacities, we first would have to develop methods in the discrete time domain that appropriately model the arising blocking effects. Given that, we could include them in the decomposition approach. Finally, discrete-time models with batch service could be applied to extend the approach by dual command cycles and vehicles with multiple load handling devices.



# Glossary of Notation

Notation	Description	Page
$a_l$	lift acceleration/deceleration rate in vertical direction (tier-to-tier configuration)	106
$a_{l_{in}}$	lift <sub>in</sub> acceleration/deceleration rate in vertical direction (tier-captive configuration)	75
$a_{l_{out}}$	lift <sub>out</sub> acceleration/deceleration rate in vertical direction (tier-captive configuration)	75
$a_{v,x}$	vehicle acceleration/deceleration rate in horizontal direction	74
$a_{v,y}$	vehicle/LHD acceleration/deceleration rate in vertical direction	74
$A$	random variable describing the inter-arrival time distribution	32
$A_j$	random variable describing the inter-arrival time distribution of queueing system $j$	55
$\alpha_{j,i}$	probability that $A_j$ assumes value $i$	55
$A_\xi$	random variable describing an inter-arrival time distribution	58
$\alpha_{\xi,i}$	probability that $A_\xi$ assumes value $i$	58
$A_\zeta$	random variable describing an inter-arrival time distribution	58
$\alpha_{\zeta,i}$	probability that $A_\zeta$ assumes value $i$	58
$A_{aisle}$	random variable describing the inter-arrival time distribution at each aisle station	114
$A_{aisle}^{(n_t)}$	random variable describing the inter-arrival time distribution at each aisle station in the $n_t$ -th iteration	124

Notation	Description	Page
$A_{aisle,R}$	random variable describing the inter-arrival time distribution of the retrieval transactions arriving at each aisle station	112
$A_{aisle,S}$	random variable describing the inter-arrival time distribution of the storage transactions arriving at each aisle station	114
$A_{aisle,S}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the storage transactions arriving at each aisle station in the $n_{it}$ -th iteration	124
$A_{lift_{in}}$	random variable describing the inter-arrival time distribution of the storage transactions at a $lift_{in}$ station	82
$A_{lift_{in}}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the storage transactions at a $lift_{in}$ station in the $n_{it}$ -th iteration	100
$A_{lift_{out}}$	random variable describing the inter-arrival time distribution of the retrieval transactions at a $lift_{out}$ station	84
$A_{lift_{out}}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the retrieval transactions at a $lift_{out}$ station in the $n_{it}$ -th iteration	100
$A_{ps,j}$	random variable describing the inter-arrival time distribution of the retrieval transactions at the $j$ -th picking station	84
$A_{ps,j}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the retrieval transactions at the $j$ -th picking station in the $n_{it}$ -th iteration	101
$A_{\bar{ps},j}$	random variable describing the inter-arrival time distribution of the transactions that are bypassing the $j$ -th picking station	84
$A_{\bar{ps},j}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the transactions that are bypassing the $j$ -th picking station in the $n_{it}$ -th iteration	101

Notation	Description	Page
$A_R$	random variable describing the inter-arrival time distribution of the retrieval transactions	81
$A_S$	random variable describing the inter-arrival time distribution of the storage transactions	81
$A_S^{(n_{it})}$	random variable describing the inter-arrival time distribution of the storage transactions in the $n_{it}$ -th iteration	100
$A_{S_{re-en}}$	random variable describing the inter-arrival time distribution of the storage transactions that re-enter the system	81
$A_{S_{re-en}}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the storage transactions that re-enter the system in the $n_{it}$ -th iteration	100
$A_{S_{rep}}$	random variable describing the inter-arrival time distribution of the storage transactions that enter the system for replenishment	81
$A_{veh}$	random variable describing the inter-arrival time distribution of the storage and retrieval transactions at a vehicle station	82
$A_{veh}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the storage and retrieval transactions at a vehicle station in the $n_{it}$ -th iteration	100
$A_{veh,R}$	random variable describing the inter-arrival time distribution of the retrieval transactions at a vehicle station	81
$A_{veh,S}$	random variable describing the inter-arrival time distribution of the storage transactions at a vehicle station	82
$A_{veh,S}^{(n_{it})}$	random variable describing the inter-arrival time distribution of the storage transactions at a vehicle station in the $n_{it}$ -th iteration	100
AGV	automated guided vehicle	19
ANOVA	analysis of variance	38
AS/RS	automated storage and retrieval system	1

Notation	Description	Page
AVS/RS	autonomous vehicle storage and retrieval system	2
$B$	random variable describing the service time distribution	32
$B_j$	random variable describing the service time distribution of queueing system $j$	55
$\beta_{j,i}$	probability that $B_j$ assumes value $i$	55
$B_{aisle}$	random variable describing the service time distribution of the aisle station	114
$\beta_{aisle,i}$	probability that $B_{aisle}$ assumes value $i$	117
$B_{aisle,R}$	random variable describing the service time distribution of the retrieval transactions of an aisle station	117
$\beta_{aisle,R,i}$	probability that $B_{aisle,R}$ assumes value $i$	117
$B_{lift_{in}}$	random variable describing the service time distribution of a lift <sub>in</sub> station	82
$\beta_{lift_{in},i}$	probability that $B_{lift_{in}}$ assumes value $i$	93
$B_{lift_{out}}$	random variable describing the service time distribution of a lift <sub>out</sub> station	84
$\beta_{lift_{out},i}$	probability that $B_{lift_{out}}$ assumes value $i$	95
$B_{ps}$	random variable describing the service time distribution of a picking station	84
$B_{veh}$	random variable describing the service time distribution of a vehicle station	82
$\beta_{veh,i}$	probability that $B_{veh}$ assumes value $i$	87
$B_{veh,R}$	random variable describing the service time distribution of the retrieval transactions of a vehicle station	88
$\beta_{veh,R,i}$	probability that $B_{veh,R}$ assumes value $i$	88
Binomial	derivation of the distribution of waiting storage	126
$(Q_j, 1 - p_R)$	transactions $Q_{j,S}$ in a queue $j$ with both storage and retrieval transactions by the use of the Binomial distribution	
$c$	index indicating a configuration	134
$C$	constant server capacity of a $G^X G^{0,C} 1 K$ queue of transportation type	52

Notation	Description	Page
$Cap$	storage capacity of the SBS/RS	152
CDF	cumulative distribution function	48
CQN	closed queueing network	33
$cv^2[X]$	squared coefficient of variation (scv) of a random variable $X$	48
$d_a$	width of an aisle including the racks on either side of the aisle	153
$d_c$	distance between two storage columns	73
$d_l$	distance between two levels of a tier	73
$d_t$	distance between two tiers	73
$d_{l_{in}}$	distance between the lowest tier and the input point of the lift <sub>in</sub>	76
$d_{l_{out}}$	distance between the lowest tier and the output point of the lift <sub>out</sub>	76
$d_{v,x}^{(x,x')}$	horizontal distance between position $x$ and $x'$	91
$d_{v,y}^{(y_l,y_{l'})}$	vertical distance between position $y_l$ and $y_{l'}$	92
$d_{l_{in}}^{(y_t,in)}$	vertical distance between position $y_t$ and the input point of the aisle	94
$d_{l_{in}}^{(in,y_t)}$	vertical distance between the input point of the aisle and position $y_t$	94
$d_{l_{out}}^{(out,y_t)}$	vertical distance between the output point of the aisle and position $y_t$	96
$D_j$	random variable describing the inter-departure time distribution of queueing system $j$	55
$\delta_{j,i}$	probability that $D_j$ assumes value $i$	55
$D_{aisle}$	random variable describing the inter-departure time distribution of the aisle station	114
$D_{aisle}^{(n_{it})}$	random variable describing the inter-departure time distribution of the aisle station in the $n_{it}$ -th iteration	124
$D_{aisle,R}$	random variable describing the inter-departure time distribution of the retrieval transaction stream of an aisle station to the picking stations	114

Notation	Description	Page
$D_{aisle,R}^{(n_{it})}$	random variable describing the inter-departure time distribution of the retrieval transaction stream of an aisle station to the picking stations in the $n_{it}$ -th iteration	124
$D_{lift_{in}}$	random variable describing the inter-departure time distribution of the storage transactions at a $lift_{in}$ station	82
$D_{lift_{in}}^{(n_{it})}$	random variable describing the inter-departure time distribution of the storage transactions at a $lift_{in}$ station in the $n_{it}$ -th iteration	100
$D_{lift_{out}}$	random variable describing the inter-departure time distribution of the retrieval transactions at a $lift_{out}$ station	84
$D_{lift_{out}}^{(n_{it})}$	random variable describing the inter-departure time distribution of the retrieval transactions at a $lift_{out}$ station in the $n_{it}$ -th iteration	101
$D_{out}$	random variable describing the inter-departure time distribution of the retrieval transactions that are leaving the system (empty bins for replenishment and bins for subsequent processes)	85
$D_{out,0.95}$	95% quantile of the inter-departure time distribution of the retrieval transactions that are leaving the system	139
$D_{out}^{(n_{it})}$	random variable describing the inter-departure time distribution of the retrieval transactions that are leaving the system in the $n_{it}$ -th iteration	101
$D_{out,j}$	random variable describing the inter-departure time distribution of the transactions before the split into the $j$ -th picking station	84
$D_{out,j}^{(n_{it})}$	random variable describing the inter-departure time distribution of the transactions before the split into the $j$ -th picking station in the $n_{it}$ -th iteration	101

Notation	Description	Page
$D_{ps,j}$	random variable describing the inter-departure time distribution of the transactions at the $j$ -th picking station	84
$D_{ps,j}^{(n_{it})}$	random variable describing the inter-departure time distribution of the transactions at the $j$ -th picking station in the $n_{it}$ -th iteration	101
$D_{veh}$	random variable describing the inter-departure time distribution of the storage and retrieval transactions at a vehicle station	82
$D_{veh}^{(n_{it})}$	random variable describing the inter-departure time distribution of the storage and retrieval transactions at a vehicle station in the $n_{it}$ -th iteration	100
$D_{veh,R}$	random variable describing the inter-departure time distribution of the retrieval transactions at a vehicle station	83
$D_{veh,R}^{(n_{it})}$	random variable describing the inter-departure time distribution of the retrieval transactions at a vehicle station in the $n_{it}$ -th iteration	100
DES	discrete-event simulation	6
Dof	degree of freedom	9
$\Delta_c^{abs}$	for configuration $c$ , absolute deviation of the value of the performance measure obtained by the decomposition approach value $v_c^{DA}$ to the value of the performance measure obtained by the discrete-event simulation value $v_c^{DES}$	134
$\Delta_c^{rel}$	for configuration $c$ , relative deviation of the value of the performance measure obtained by the decomposition approach value $v_c^{DA}$ to the value of the performance measure obtained by the discrete-event simulation value $v_c^{DES}$	134
$ \Delta^{abs} _{avg}$	average absolute deviations of $\Delta_c^{abs}$	134
$ \Delta^{rel} _{avg}$	average absolute deviations of $\Delta_c^{rel}$	134

Notation	Description	Page
$\bar{\Delta}_{cv^2[X],c}^{rel}$	for configuration $c$ , average relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the variability of random variable $X$	163
$\bar{\Delta}_{cv^2[X],all}^{rel}$	given all configurations $n_{config}^{nev}$ , overall average relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the variability of random variable $X$	164
$\Delta_{cv^2[X],c}^{rel,max}$	for configuration $c$ , maximum relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the variability of random variable $X$	163
$\Delta_{cv^2[X],all}^{rel,max}$	given all configurations $n_{config}^{nev}$ , overall maximum relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the variability of random variable $X$	164
$\Delta_{cv^2[X],c}^{rel,min}$	for configuration $c$ , minimum relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the variability of random variable $X$	163
$\Delta_{cv^2[X],all}^{rel,min}$	given all configurations $n_{config}^{nev}$ , overall minimum relative reduction of the 95% quantile of the retrieval transaction time caused by the reduction of the variability of random variable $X$	164
$\bar{\Delta}_{noRe-en,c}^{rel}$	for configuration $c$ , average relative reduction of the 95% quantile of the retrieval transaction time caused by no re-entrance of the bins	165
$\bar{\Delta}_{noRe-en,all}^{rel}$	given all configurations $n_{config}^{nev}$ , overall average relative reduction of the 95% quantile of the retrieval transaction time caused by no re-entrance of the bins	165
$\Delta_{noRe-en,c}^{rel,max}$	for configuration $c$ , maximum relative reduction of the 95% quantile of the retrieval transaction time caused by no re-entrance of the bins	165

Notation	Description	Page
$\Delta_{noRe-en,all}^{rel,max}$	given all configurations $n_{config}^{nev}$ , overall maximum relative reduction of the 95% quantile of the retrieval transaction time caused by no re-entrance of the bins	165
$\Delta_{noRe-en,c}^{rel,min}$	for configuration $c$ , minimum relative reduction of the 95% quantile of the retrieval transaction time caused by no re-entrance of the bins	165
$\Delta_{noRe-en,all}^{rel,min}$	given all configurations $n_{config}^{nev}$ , overall minimum relative reduction of the 95% quantile of the retrieval transaction time caused by no re-entrance of the bins	165
$E[X]$	expectation of a random variable $X$	48
$E[X^2]$	2nd moment of a random variable $X$	48
$\varepsilon$	value of the $\varepsilon$ -environment of the algorithms of Grassmann and Jain (1989)	49
$\varepsilon_{Split}$	threshold value of split algorithms	61
$\varepsilon_{it}$	threshold value of iterative algorithms to compute the network performance measures	101
FFT	Fast Fourier Transformation	62
FSA	fast split approximation	62
FCFS	first-come-first-serve	25
FJQN	fork-join queueing network	43
Footprint	footprint of the SBS/RS without the lifts, buffers and the conveyor system in front of the aisles	155
$G$	process distributed according to a general distribution	32
$G G 1(A_j, B_j)$	notation for the use of one of the methods to generate $W_j$ , $D_j$ or $Q_j$ of queueing system $j$	102
$\gamma^*(k, \theta)$	Gamma distribution with shape parameter $k$ and scale parameter $\theta$ , which is discretized according to the constant time increment $t_{inc}$	63
$h_x$	horizontal position of the vehicle/LHD within the tier when being requested	87
$h_y$	vertical position of the vehicle/LHD within the tier when being requested	87

Notation	Description	Page
$h_{yt}$	vertical position of the lift/tier-to-tier vehicle when being requested	93
$H_a$	tuple which is describing the horizontal position at column $h_x$ as well as the vertical position at tier $h_{yt}$ and level $h_{yl}$ of the tier-to-tier vehicle/LHD within the aisle when being requested	117
$H_{l_{in}}$	variable which is describing the vertical position $h_{yt}$ of the incoming lift when being requested	93
$H_{l_{out}}$	variable which is describing the vertical position $h_{yt}$ of the outgoing lift when being requested	95
$H_v$	tuple which is describing the horizontal position at column $h_x$ and the vertical position at level $h_{yl}$ of the tier-captive vehicle/LHD within the tier when being requested	87
$\bar{h}_x$	horizontal position of the storage/retrieval location of the transaction to store/retrieve the goods	87
$\bar{h}_{yl}$	target level of the transaction to store/retrieve the goods	87
$\bar{h}_{yt}$	target tier of the transaction to store/retrieve the goods	93
$\bar{H}_a$	tuple which is describing the horizontal position at column $\bar{h}_x$ as well as the vertical position at tier $\bar{h}_{yt}$ and level $\bar{h}_{yl}$ of the tier-to-tier vehicle/LHD to store/retrieve the goods	117
$\bar{H}_{l_{in}}$	variable which is describing the target tier $\bar{h}_{yt}$ of the lift <sub>in</sub>	93
$\bar{H}_{l_{out}}$	variable which is describing the position of the tier $\bar{h}_{yt}$ from where the bin is being retrieved	95
$\bar{H}_v$	tuple which is describing the horizontal position at column $\bar{h}_x$ and the vertical position at level $\bar{h}_{yl}$ of the tier-captive vehicle/LHD to store/retrieve the goods	87
$i$	index indicating time increment or number of customer	47
$in$	vertical position of the input point of the lift	92

Notation	Description	Page
i.i.d.	independent and identically distributed	50
I/O	input/output	9
IQM	integrated queueing model consisting of semi-open queueing networks and additional queueing systems	32
$j$	index indicating queueing system $j$	55
$k$	shape parameter of Gamma distribution	63
$K$	constant batch size for collection and service of a $G^X G^{K,K} 1$ queueing system (only used in literature review)	51
$K$	maximum number of customers in a queueing system with finite buffer capacity (only used in literature review)	42
$l$	index	48
$L$	threshold value of a $G^X G^{L,K} 1$ queueing system	51
LHD	load handling device	9
LSM	load shuffling model	40
$\lambda_R$	arrival rate of the retrieval transactions	81
$\lambda_S$	arrival rate of the storage transactions	81
$\lambda_{S_{rep}}$	arrival rate of the storage transactions for replenishment	81
$\lambda_{S_{re-en}}$	arrival rate of the re-entering storage transactions	81
$m$	number of servers	32
$M$	Markov process	32
$max_H$	maximum height of the system	152
$max_L$	maximum length of the system	152
$max_W$	maximum width of the system	152
$\text{Merge}(A_\xi, A_\zeta)$	notation for the use of the merging operation to merge two different distributions $A_\xi$ and $A_\zeta$	102
$\text{Merge}(A_\xi, n_{\text{Merge}}\text{-times})$	notation for the use of the merging operation to merge $n_{\text{Merge}}$ identical distributions $A_\xi$	102
MGMM	matrix geometric method	34
ML	multi-level shuttle	2
MOOM	multi-objective optimization model with integrated cycle time model	40

Notation	Description	Page
MPA	manufacturing performance analyzer	36
$\mu(v)$	service rate $\mu$ which is dependent on the number of customers being served $v$ (only used in literature review)	34
$\mu$	service rate	34
$n$	index	62
$n_a$	number of aisles	73
$n_{a,max}$	maximum number of aisles	153
$n_{a,min}(n_l)$	minimum number of aisles needed to reach the storage capacity for a given number of levels per tier $n_l$	154
$n_c$	number of columns on either side of an aisle	73
$n_{c,max}$	maximum number of columns on either side of an aisle	154
$n_c(n_a, n_t, n_l)$	number of columns needed on either side of the aisle such that the required capacity for a given number of aisles $n_a$ , tiers $n_t$ , and levels per tier $n_l$ is reached	155
$n_{config}^{val}$	number of analyzed configurations in the validation section	134
$n_{config}^{nev}$	number of analyzed configurations in the numerical evaluation	164
$n_{FSA}$	threshold value of the fast split approximation	62
$n_{it}$	number of iterations in the algorithms to compute the network performance measures	99
$n_l$	number of levels per tier	73
$n_{L,tot}$	total number of lifts in the system	155
$n_{mix}$	number of distributions of a mixture distribution	64
$n_{Merge}$	number of identical distributions that are being merged	102
$n_{ps}$	number of picking stations	76
$n_{sum}$	number of convolution operations that are needed to reach the threshold value of the split operations	61
$n_t$	number of tiers per aisle	73
$n_{t,max}$	maximum number of tiers per aisle	154

Notation	Description	Page
$n_{t,min}(n_a, n_l)$	minimum number of tiers per aisle needed to reach the storage capacity for a given number of aisles $n_a$ and levels per tier $n_l$	154
$n_V$	number of vehicles in the system	155
$N_j$	random variable describing the distribution of the number of customers in queueing system $j$ at the arrival instant of a customer	57
$\eta_{j,i}$	probability that $N_j$ assumes value $i$	57
$\mathcal{N}^*(\bar{x}, \sigma^2)$	normal distribution with mean $\bar{x}$ and variance $\sigma^2$ , which is discretized according to the constant time increment $t_{inc}$	62
$out$	vertical position of the output point of the lift	94
OQN	open queueing network	36
$p_{em}$	probability that a bin is empty after picking	76
$p_{ps}$	probability that the target of the retrieval transaction is a picking station	76
$p_R$	ratio of the retrieval transactions to the sum of storage and retrieval transactions	82
$p_\xi$	probability that a customer is routed to direction $\xi$	102
pmf	probability mass function	47
$P$	probability measure	47
$P[TT, H_a, \tilde{H}_a]_a$	probability of an arbitrary transaction cycle of the tier-to-tier vehicle (aisle station)	117
$P[TT, H_v, \tilde{H}_v]_v$	probability of an arbitrary transaction cycle of the tier-captive vehicle	87
$P[S, H_{l_{in}}, \tilde{H}_{l_{in}}]_{l_{in}}$	probability of an arbitrary transaction cycle of the incoming lift	93
$P[R, out, \tilde{H}_{l_{out}}]_{l_{out}}$	probability of an arbitrary transaction cycle of the outgoing lift	95
$PH$	process that is distributed according to a phase-type distribution	34
POSC	point of service completion	27
$\Phi$	random variable	49
$\phi_l$	probability that $\Phi$ assumes value $l$	49

---

Notation	Description	Page
$Q_j$	random variable describing the distribution of the number of waiting customers at the arrival instant of a customer at queueing system $j$	98
$q_{j,i}$	probability that there are $i$ customers waiting at the arrival instant of a customer at queueing system $j$	98
$Q_{j,S}$	random variable describing the distribution of the number of waiting storage transactions at the arrival instant of a customer at queueing system $j$	123
$q_{j,S,i}$	probability that there are $i$ storage transactions waiting at the arrival instant of a customer at queueing system $j$	123
$Q_{aisle}$	random variable describing the queue length distribution of an aisle station at the arrival instant of a customer	124
$Q_{aisle,S}$	random variable describing the storage transaction queue length distribution of an aisle station at the arrival instant of a customer	121
$Q_{aisle,S,0.95}$	95% quantile of the storage transaction queue length distribution of an aisle station at the arrival instant of a customer	148
$Q_{lift_{in}}$	random variable describing the queue length distribution of a lift <sub>in</sub> station at the arrival instant of a customer	96
$Q_{lift_{in},0.95}$	95% quantile of the queue length distribution of a lift <sub>in</sub> station at the arrival instant of a customer	140
$Q_{ps,j}$	random variable describing the queue length distribution of the $j$ -th picking station at the arrival instant of a customer	96
$Q_{ps,j,0.95}$	95% quantile of the queue length distribution of the $j$ -th picking station at the arrival instant of a customer	141
QNA	queueing network analyzer	36

Notation	Description	Page
$R$	variable indicating that the transaction is of type retrieval	87
RIO	return to I/O	27
$[R, out, \bar{H}_{l_{out}}]_{l_{out}}$	triple that classifies an arbitrary transaction cycle of the lift <sub>out</sub> station	95
$\rho_{aisle}$	utilization of an aisle station	121
$\rho_{veh}$	utilization of a vehicle	96
$\rho_{lift_{in}}$	utilization of an incoming lift	96
$\rho_{lift_{out}}$	utilization of an outgoing lift	96
$\rho_{max}$	maximum utilization of both the vehicles and the lifts in the numerical evaluation	152
$\rho_{max,l}$	maximum utilization of the lifts	155
$\rho_{ps}$	utilization of a picker / picking station	96
$S$	variable indicating that the transaction is of type storage	87
SBCSS	shuttle-based compact storage system	36
SBS/RS	shuttle-based storage and retrieval system	2
scv	squared coefficient of variation	48
SL	single-level shuttle	2
SOQN	semi-open queueing network	34
SP-AS/RS	split-platform automated storage and retrieval system	41
$\text{Split}(A_j, p_\xi)$	notation for the use of the split operation to determine the distribution $A_\xi$ after the split of the distribution $A_j$ in two or more directions	102
S/R	storage/retrieval	38
SRM	storage and retrieval machine	1
$[S, H_{l_{in}}, \bar{H}_{l_{in}}]_{l_{in}}$	triple that classifies an arbitrary transaction cycle of the lift <sub>in</sub> station	93
$\sigma^2$	variance of a random variable	62
$\hat{\sigma}^2$	variance of a mixture distribution	63

Notation	Description	Page
$t_{inc}$	constant time increment	47
$t_l^{(yt,yt')}$	travel time of the lift between any given positions $yt$ and $yt'$	118
$t_{l_{in},trans}$	lift <sub>in</sub> transfer time, i.e. time to load or unload the goods onto or from the incoming lift	75
$t_{l_{in}}^{(yt,yt')}$	travel time of the incoming lift between any given positions $yt$ and $yt'$	93
$t_{l_{out},trans}$	lift <sub>out</sub> transfer time, i.e. time to load or unload the goods onto or from the outgoing lift	76
$t_{l_{out}}^{(yt,yt')}$	travel time of the outgoing lift between any given positions $yt$ and $yt'$	95
$t_{v,trans}$	vehicle transfer time, i.e. time to load or unload the goods onto or from the vehicle	74
$t_v^{(x,yl),(x',yl')}$	travel time of the vehicle between any given positions $(x,yl)$ and $(x',yl')$	88
$t_{v,x}^{(x,x')}$	horizontal movement time from position $x$ to $x'$	91
$t_{v,y}^{(yl,yl')}$	vertical movement time from position $yl$ to $yl'$	91
$t[TT,H_a,\bar{H}_a]_a$	service time of an arbitrary transaction cycle of the tier-to-tier vehicle (aisle station)	117
$t[TT,H_v,\bar{H}_v]_v$	service time of an arbitrary transaction cycle of the tier-captive vehicle	87
$t[R,out,\bar{H}_{l_{out}}]_{l_{out}}$	service time of an arbitrary transaction cycle of the outgoing lift	95
$t[S,H_{l_{in}},\bar{H}_{l_{in}}]_{l_{in}}$	service time of an arbitrary transaction cycle of the incoming lift	93
$T_j$	random variable describing the sojourn time distribution of queueing system $j$	55
$\tau_{j,i}$	probability that $T_j$ assumes value $i$	55
$T_R$	random variable describing the retrieval transaction time distribution	96
$T_{R,0.95}$	95% quantile of the retrieval transaction time distribution	138

Notation	Description	Page
$T_R^{(n_{it})}$	random variable describing the retrieval transaction time distribution in the $n_{it}$ -th iteration	101
$T_{veh,R}$	random variable describing the sojourn time distribution of the retrieval transactions at a vehicle station	98
$T_{veh,R}^{(n_{it})}$	random variable describing the sojourn time distribution of the retrieval transactions at a vehicle station in the $n_{it}$ -th iteration	100
$T_{lift_{out}}$	random variable describing the sojourn time distribution of the retrieval transactions at a lift <sub>out</sub> station	98
$T_{lift_{out}}^{(n_{it})}$	random variable describing the sojourn time distribution of the retrieval transactions at a lift <sub>out</sub> station in the $n_{it}$ -th iteration	100
TT	variable that classifies the type of transaction	87
$[TT, H_v, \bar{H}_v]_v$	triple that classifies an arbitrary transaction cycle of the tier-captive vehicle	87
$[TT, H_a, \bar{H}_a]_a$	triple that classifies an arbitrary transaction cycle of the tier-to-tier vehicle (aisle station)	117
$\theta$	scale parameter of a Gamma distribution	63
$u$	index for quantile value	48
$v_l$	lift velocity in vertical direction (tier-to-tier configuration)	106
$v_{l,in}$	lift <sub>in</sub> velocity in vertical direction	75
$v_{l,out}$	lift <sub>out</sub> velocity in vertical direction	75
$v_{v,x}$	vehicle velocity in horizontal direction	74
$v_{v,y}$	vehicle/LHD velocity in vertical direction	74
$\text{VAR}[X]$	variance of a random variable $X$	48
$\text{value}^{DA}$	value of the performance measure obtained by the decomposition approach	134
$\text{value}^{DES}$	value of the ensemble average of the performance measure obtained by the discrete-event simulation	134
VDI	Association of German Engineers	41

Notation	Description	Page
$w_l$	weights of a mixture distribution	64
$w_{sum}$	weighted sum that is used for the computations of the fast split approximation	63
$W_j$	random variable describing the waiting time distribution of queueing system $j$	55
$\omega_{j,i}$	probability that $W_j$ assumes value $i$	55
$W_{aisle}$	random variable describing the waiting time distribution of the storage and retrieval transactions at an aisle station	123
$W_{aisle}^{(n_{it})}$	random variable describing the waiting time distribution of the storage and retrieval transactions at an aisle station in the $n_{it}$ -th iteration	124
$W_{lift_{out}}$	random variable describing the waiting time distribution of the retrieval transactions at a $lift_{out}$ station	98
$W_{lift_{out}}^{(n_{it})}$	random variable describing the waiting time distribution of the retrieval transactions at a $lift_{out}$ station in the $n_{it}$ -th iteration	100
$W_{veh}$	random variable describing the waiting time distribution of the storage and retrieval transactions at a vehicle station	98
$W_{veh}^{(n_{it})}$	random variable describing the waiting time distribution of the storage and retrieval transactions at a vehicle station in the $n_{it}$ -th iteration	100
$x$	possible horizontal position of the vehicle/LHD within the tier	86
$(x, yl)$	tuple which is describing the horizontal position at column $x$ and the vertical position at level $yl$ of the tier-captive vehicle/LHD within the tier	86
$(x, yt, yl)$	triple which is describing the horizontal position at column $x$ as well as the vertical position at tier $yt$ and level $yl$ of the tier-to-tier vehicle/LHD within the aisle	116

Notation	Description	Page
$X$	random variable describing the batch size distribution of a queueing system with batch arrivals (only used in literature review)	51
$X$	arbitrary random variable	47
$\chi_i$	probability that random variable $X$ assumes value $i$	47
$X_u$	$u\%$ quantile of a discrete random variable $X$	48
$\bar{x}$	mean of a random variable	62
$\hat{x}$	mean of a mixture distribution	63
$\xi$	notation of node $\xi$	59
$yl$	possible vertical position of the vehicle/LHD within the tier	86
$yt$	possible vertical position of the incoming/outgoing lift and the tier-to-tier vehicle/LHD within the aisle	92
$Y$	arbitrary random variable	49
$y_i$	probability that random variable $Y$ assumes value $i$	49
$Z$	routing matrix	81
$z_{j,\xi}$	routing probability from node $j$ to $\xi$	59
$z_{j,\zeta}$	routing probability from node $j$ to $\zeta$	59
$\zeta$	notation of node $\zeta$	59
$\otimes$	convolution operator	49



# Bibliography

- Arnold, D. and K. Furmans (2009). *Materialfluss in Logistiksystemen* (6 ed.). Springer Berlin Heidelberg.
- Bolch, G., S. Greiner, H. de Meer and K. S. Trivedi (1998). *Queueing Networks and Markov Chains: modeling and performance evaluation with computer science applications*. New York: John Wiley & Sons, Inc.
- Borovinsek, M., B. Y. Ekren, A. Burinskiene and T. Lerher (2017). Multi-objective optimisation model of shuttle-based storage and retrieval system. *Transport* 32(2), p. 120–137.
- Buitenhok, R., G.-J. van Houtum and H. Zijm (2000). AMVA-based solution procedures for open queueing networks with population constraints. *Annals of Operations Research* 93, p. 15–40.
- Cai, X., S. S. Heragu and Y. Liu (2014). Modeling and evaluating the AVS/RS with tier-to-tier vehicles using a semi-open queueing network. *IIE Transactions* 46(9), p. 905–927.
- Deb, K., A. Pratap, S. Agarwal and T. Meyarivan (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2), p. 182–197.
- Eder, M. and G. Kartnig (2016a). Durchsatzoptimierung von Shuttle-Systemen mithilfe eines analytischen Berechnungsmodells. In: *Logistics Journal Proceedings*, p. 171–177.
- Eder, M. and G. Kartnig (2016b). Throughput Analysis of S/R Shuttle Systems and Ideal Geometry for High Performance. *FME Transactions* 44(2), p. 174–179.

- Ekren, B. Y. (2017). Graph-based solution for performance evaluation of shuttle-based storage and retrieval system. *International Journal of Production Research* 55(21), p. 6516–6526.
- Ekren, B. Y. and S. S. Heragu (2010a). Approximate analysis of load-dependent generally distributed queuing networks with low service time variability. *European Journal of Operational Research* 205, p. 381–389.
- Ekren, B. Y. and S. S. Heragu (2010b). Simulation-based regression analysis for the rack configuration of an autonomous vehicle storage and retrieval system. *International Journal of Production Research* 48(21), p. 6257–6274.
- Ekren, B. Y. and S. S. Heragu (2011). Simulation based performance analysis of an autonomous vehicle storage and retrieval system. *Simulation Modelling Practice and Theory* 19, p. 1640–1650.
- Ekren, B. Y. and S. S. Heragu (2012). Performance comparison of two material handling systems: AVS/RS and CBAS/RS. *International Journal of Production Research* 50(15), p. 4061–4074.
- Ekren, B. Y., S. S. Heragu, A. Krishnamurthy and C. J. Malmborg (2013). An Approximate Solution for Semi-Open Queueing Network Model of an Autonomous Vehicle Storage and Retrieval System. *IEEE Transactions on Automation Science and Engineering* 10(1), p. 205–215.
- Ekren, B. Y., S. S. Heragu, A. Krishnamurthy and C. J. Malmborg (2014). Matrix-geometric solution for semi-open queueing network model of autonomous vehicle storage and retrieval system. *Computers & Industrial Engineering* 68, p. 78–86.
- Ekren, B. Y., Z. Sari and T. Lerher (2015). Warehouse Design under Class-Based Storage Policy of Shuttle-Based Storage and Retrieval System. In: *15th IFAC Symposium on Information Control Problems in Manufacturing*, Volume 48-3, p. 1152–1154.
- Epp, M., P. Pagani, J. Stoll, S. Scherer, C. Rohlehr and K. Furmans (2016). Performance evaluation of closed-loop logistics systems with generally distributed service times. In: *Proceedings of the Karlsruhe Service Summit Research Workshop 2016, Karlsruhe, Germany*.

- Epp, M., J. Stoll, S. Scherer, P. Pagani and K. Furmans (2015). A decomposition approach for the calculation of the cycle time distribution of closed queueing systems. In: *10th conference on stochastic models of manufacturing and service operations, Volos, Greece*.
- Epp, M., S. Wiedemann and K. Furmans (2017). A discrete-time queueing network approach to performance evaluation of autonomous vehicle storage and retrieval systems. *International Journal of Production Research* 55(4), p. 960–978.
- Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer New York.
- Fukunari, M., K. Bennett and C. Malmborg (2004). Decision-tree learning in dwell point policies in autonomous vehicle storage and retrieval systems (AVSRS). In: *International Conference on Machine Learning and Applications, Louisville, USA*, p. 81–84.
- Fukunari, M. and C. J. Malmborg (2008). An efficient cycle time model for autonomous vehicle storage and retrieval systems. *International Journal of Production Research* 46(12), p. 3167–3184.
- Fukunari, M. and C. J. Malmborg (2009). A network queuing approach for evaluation of performance measures in autonomous vehicle storage and retrieval systems. *European Journal of Operational Research* 193(1), p. 152–167.
- Furmans, K. (2000). *Bedientheoretische Methoden als Hilfsmittel der Materialflußplanung*. Number 52 in Wissenschaftliche Berichte des Institutes für Fördertechnik und Logistiksysteme des Karlsruher Instituts für Technologie. Institut für Fördertechnik und Logistiksysteme.
- Furmans, K. (2004). A Framework of Stochastic Finite Elements for Models of Material Handling Systems. In: *Progress in Material Handling Research*.
- Furmans, K., T. Schmidt, I. Meinhardt, M. Epp and M. Rimmele (2015). Analytical computation of sojourn time distributions in large-scale conveyer systems. In: *10th conference on stochastic models of manufacturing and service operations, Volos, Greece*.

- Furmans, K. and A. Zillus (1996). Modeling Independent Production Buffers in Discrete Time Queueing Networks. In: *Proceedings of CIMAT '96, Grenoble, France*, p. 275–280.
- Grassmann, W. K. and J. L. Jain (1989). Numerical Solutions of the Waiting Time Distribution and Idle Time Distribution of the Arithmetic  $GI/G/1$  Queue. *Operations Research* 37(1), p. 141–150.
- Heragu, S. S., X. Cai, A. Krishnamurthy and C. J. Malmborg (2011). Analytical models for analysis of automated warehouse material handling systems. *International Journal of Production Research* 49(22), p. 6833–6861.
- Hu, Y.-H., S. Y. Huang, C. Chen, W.-J. Hsu, A. C. Toh, C. K. Loh and T. Song (2005). Travel time analysis of a new automated storage and retrieval system. *Computers & Operations Research* 32(6), p. 1515 – 1544.
- Hu, Y.-H., Z. D. Zhu and W.-J. Hsu (2010). Load shuffling algorithms for split-platform AS/RS. *Robotics and Computer-Integrated Manufacturing* 26, p. 677–685.
- Jain, J. L. and W. K. Grassmann (1988). Numerical solution for the departure process from the  $GI/G/1$  queue. *Computers & Operations Research* 15(3), p. 293–296.
- Kartnig, G. and J. Oser (2014). Throughput analysis of S/R shuttle systems. In: *Progress in Material Handling Research*.
- Krämer, W. and M. Langenbach-Belz (1976). Approximate formulae for the delay in the queueing system  $GI/G/1$ . In: *8th International Teletraffic Congress, Melbourne, Australia*, p. 235.1–235.8.
- Krishnamurthy, A., D. Roy, S. S. Heragu and C. J. Malmborg (2010). Blocking Effects on Performance of Warehouse Systems with Autonomous Vehicles. In: *Progress in Material Handling Research*.
- Kumar, A., D. Roy and M. Tiwari (2014). Optimal partitioning of vertical zones in vehicle-based warehouse systems. *International Journal of Production Research* 52(5), p. 1285–1305.

- Kuo, P. H., A. Krishnamurthy and C. J. Malmborg (2007). Design models for unit load storage and retrieval systems using autonomous vehicle technology and resource conserving storage and dwell point policies. *Applied Mathematical Modelling* 31(10), p. 2332–2346.
- Kuo, P. H., A. Krishnamurthy and C. J. Malmborg (2008). Performance modelling of autonomous vehicle storage and retrieval systems using class-based storage policies. *International Journal of Computer Applications in Technology* 31(3/4), p. 238–248.
- Lerher, T. (2016a). Multi-tier Shuttle-based Storage and Retrieval Systems. *FME Transactions* 44, p. 285–290.
- Lerher, T. (2016b). *Throughput and Energy Related Performance Calculations for Shuttle Based Storage and Retrieval Systems*. Nova Science Publishers, New York.
- Lerher, T. (2016c). Travel time model for double-deep shuttle-based storage and retrieval systems. *International Journal of Production Research* 54(9), p. 2519–2540.
- Lerher, T., B. Y. Eken, G. Dukic and B. Rosi (2015). Travel time model for shuttle-based storage and retrieval systems. *The International Journal of Advanced Manufacturing Technology* 78(9), p. 1705–1725.
- Lerher, T., B. Y. Eken, Z. Sari and B. Rosi (2015). Simulation Analysis of Shuttle Based Storage and Retrieval Systems. *International Journal of Simulation Modelling* 14(1), p. 48–59.
- Liekenbrock, D. (2016). Shuttlesysteme und Klein-RBG. In: *VDI-Berichte Nr. 2275*.
- Malmborg, C. J. (2002). Conceptualizing tools for autonomous vehicle storage and retrieval systems. *International Journal of Production Research* 40(8), p. 1807–1822.
- Malmborg, C. J. (2003). Interleaving dynamics in autonomous vehicle storage and retrieval systems. *International Journal of Production Research* 41(5), p. 1057–1069.

- Marchet, G., M. Melacini, S. Perotti and E. Tappia (2012). Analytical model to estimate performances of autonomous vehicle storage and retrieval systems for product totes. *International Journal of Production Research* 50(24), p. 7134–7148.
- Marchet, G., M. Melacini, S. Perotti and E. Tappia (2013). Development of a framework for the design of autonomous vehicle storage and retrieval systems. *International Journal of Production Research* 51(14), p. 4365–4387.
- Matzka, J. (2011). *Discrete Time Analysis of Multi-Server Queueing Systems in Material Handling and Service*. Ph.D. thesis, Karlsruhe Institute of Technology.
- Meng, G. and S. S. Heragu (2004). Batch size modeling in a multi-item, discrete manufacturing system via an open queuing network. *IIE Transactions* 36(8), p. 743–753.
- Ning, Z., L. Lei, Z. Saipeng and G. Lodewijks (2016). An efficient simulation model for rack design in multi-elevator shuttle-based storage and retrieval system. *Simulation Modelling Practice and Theory* 67, p. 100–116.
- Nussbaumer, H. J. (1982). *Fast Fourier Transform and Convolution Algorithms*. Springer Berlin Heidelberg.
- Özden, E. (2011). *Discrete time Analysis of Consolidated Transport Processes*. Ph.D. thesis, Karlsruhe Institute of Technology.
- Özden, E. and K. Furmans (2010). Analysis of the discrete-time  $G^X|G^{[L,K]}|1$ -queue. In: *24th European Conference on Operational Research, Lisbon, Portugal*.
- Roodbergen, K. J. and I. F. A. Vis (2009). A survey of literature on automated storage and retrieval systems. *European Journal of Operational Research* 194, p. 343–362.
- Roy, D. and A. Krishnamurthy (2011). Improving Throughput Capacity in Multi-tier Warehouses with Autonomous Vehicles. In: *Industrial Engineering Research Conference, Reno, USA*.

- Roy, D., A. Krishnamurthy, S. S. Heragu and C. J. Malmborg (2012). Performance analysis and design trade-offs in warehouses with autonomous vehicle technology. *IIE Transactions* 44(12), p. 1045–1060.
- Roy, D., A. Krishnamurthy, S. S. Heragu and C. J. Malmborg (2014). Blocking Effects in Warehouse Systems with Autonomous Vehicles. *IEEE Transactions on Automation Science and Engineering* 11(2), p. 439–451.
- Roy, D., A. Krishnamurthy, S. S. Heragu and C. J. Malmborg (2015a). Queuing models to analyze dwell-point and cross-aisle location in autonomous vehicle-based warehouse systems. *European Journal of Operational Research* 242, p. 72–87.
- Roy, D., A. Krishnamurthy, S. S. Heragu and C. J. Malmborg (2015b). Stochastic Models for Unit-Load Operations in Warehouse Systems with Autonomous Vehicles. *Annals of Operations Research* 231, p. 129–155.
- Sari, Z., L. Ghomri, B. Y. Eken and T. Lerher (2014). Experimental validation of travel time models for shuttle-based automated storage and retrieval system. In: *Progress in Material Handling Research*.
- Schleyer, M. (2007). *Discrete Time Analysis of Batch Processes in Material Flow Systems*. Ph.D. thesis, Universität Karlsruhe (TH).
- Schleyer, M. and K. Furmans (2007). An analytical method for the calculation of the waiting time distribution of a discrete time G/G/1-queueing system with batch arrivals. *OR Spectrum* 29(4), p. 745–763.
- Schmidt, T. (2010). Shuttle gegen AKL: Konkurrenten und Weggefährten. *LOGISTRA* 7-8, p. 12–15.
- Schwarz, J. A. and M. Epp (2016). Performance evaluation of a transportation-type bulk queue with generally distributed inter-arrival times. *International Journal of Production Research* 54(20), p. 6251–6264.
- Tappia, E., D. Roy, R. de Koster and M. Melacini (2017). Modeling, Analysis, and Design Insights for Shuttle-based Compact Storage Systems. *Transportation Science* 51(1), p. 269–295.

- ten Hompel, M., T. Schmidt and L. Nagel (2007). *Materialflusssysteme: Förder- und Lagertechnik* (3 ed.). Springer Berlin Heidelberg.
- Tran-Gia, P. (1996). *Analytische Leistungsbewertung verteilter Systeme: Eine Einführung*. Springer Berlin Heidelberg.
- VDI-Richtlinie 2692 (2015). *Shuttle-Systeme für kleine Ladeeinheiten*. Düsseldorf: VDI-Verlag GmbH. VDI-Gesellschaft Materialfluß Fördertechnik Logistik (Hrsg.).
- Whitt, W. (1983). The Queueing Network Analyzer. *The Bell System Technical Journal* 62(9), p. 2779–2815.
- Zhang, L., A. Krishnamurthy, C. J. Malmborg and S. S. Heragu (2009). Variance-based approximations of transaction waiting times in autonomous vehicle storage and retrieval systems. *European Journal of Industrial Engineering* 3(2), p. 146–169.
- Zou, B., X. Xu, Y. Gong and R. de Koster (2016). Modeling parallel movement of lifts and vehicles in tier-captive vehicle-based warehousing systems. *European Journal of Operational Research* 524, p. 51–67.

# List of Figures

1.1	Outline of the thesis . . . . .	7
2.1	Position of the SBS/RS within the material flow of a warehouse . . . . .	10
2.2	Different notations in literature based on the DoF of the vehicles . . . . .	11
2.3	SBS/RS design (based on the AS/RS design of Roodbergen and Vis (2009)) . . . . .	12
2.4	System load – flow of storage and retrieval transactions . . . . .	13
2.5	Example of an SBS/RS with a conveyor loop and two picking stations . . . . .	16
2.6	Physical design . . . . .	17
2.7	Control policies (extension of the AS/RS policies of Roodbergen and Vis (2009)) . . . . .	23
2.8	Design requirements and performance measures (based on the AS/RS requirements and measures of Roodbergen and Vis (2009)) . . . . .	28
4.1	A selection of discrete-time models and algorithms . . . . .	51
4.2	An example of a closed queueing network . . . . .	52
4.3	An example of an open queueing network . . . . .	53
4.4	$G G 1$ model of a queueing system $j$ . . . . .	56
4.5	Model of a stochastic merge . . . . .	58
4.6	Model of a stochastic split . . . . .	60
5.1	Tier-captive system with 3 aisles and 2 picking stations . . . . .	74
5.2	Plan view of the tier-captive system . . . . .	75
5.3	Aisle view of the tier-captive system . . . . .	76
5.4	Graph model of the tier-captive system . . . . .	78
5.5	Queueing model of the tier-captive system . . . . .	83
5.6	Tier-to-tier system with 3 aisles and 2 picking stations . . . . .	105
5.7	Plan view of the tier-to-tier system . . . . .	106

5.8	Aisle view of the tier-to-tier system . . . . .	107
5.9	Graph model of the tier-to-tier system . . . . .	109
5.10	Queueing model of the tier-to-tier system . . . . .	113
6.1	Cumulated distribution of the relative deviations of the retrieval transaction time $T_R$ regarding the expected value $E[T_R]$ and the 95% quantile $T_{R,0.95}$ given the stochastic split and the fast split methods . . . . .	138
6.2	Cumulated distribution of the relative deviations of the inter-departure time of the transactions leaving the system $D_{out}$ regarding the expected value $E[D_{out}]$ and the 95% quantile $D_{out,0.95}$ given the stochastic split and the fast split methods . . . . .	140
6.3	Cumulated distribution of the absolute deviations of the number of bins waiting in front of an incoming lift station at the arrival instant $Q_{lift_m}$ regarding the expected value $E[Q_{lift_m}]$ and the 95% quantile $Q_{lift_m,0.95}$ given the stochastic split and the fast split methods . . . . .	141
6.4	Cumulated distribution of the absolute deviations of the number of bins waiting in front of the first picking station $Q_{ps,1}$ regarding the expected value $E[Q_{ps,1}]$ and the 95% quantile $Q_{ps,1,0.95}$ given the stochastic split and the fast split methods . . . . .	142
6.5	Comparison of the probability distributions of the retrieval transaction time $T_R$ obtained by the decomposition approach using the stochastic split method and the ensemble average ( $\emptyset$ ) of 10 simulation runs for one of the tier-captive configurations . . . . .	144
6.6	Cumulated distribution of the relative deviations of the retrieval transaction time $T_R$ regarding the expected value $E[T_R]$ and the 95% quantile $T_{R,0.95}$ given the stochastic split and the fast split methods . . . . .	146
6.7	Cumulated distribution of the relative deviations of the inter-departure time of the transactions leaving the system $D_{out}$ regarding the expected value $E[D_{out}]$ and the 95% quantile $D_{out,0.95}$ given the stochastic split and the fast split methods . . . . .	147
6.8	Cumulated distribution of the absolute deviations of the number of bins waiting in front of an aisle station at the arrival instant $Q_{aisle,S}$ regarding the expected value $E[Q_{aisle,S}]$ and the 95% quantile $Q_{aisle,S,0.95}$ given the stochastic split and the fast split methods . . . . .	148

6.9 Cumulated distribution of the absolute deviations of the number of bins waiting in front of the first picking station $Q_{ps,1}$ regarding the expected value $E[Q_{ps,1}]$ and the 95% quantile $Q_{ps,1,0.95}$ given the stochastic split and the fast split methods . . . . .	149
--	-----



# List of Tables

3.1	Existing analytical models – Literature review on AVS/RSs . . . . .	32
3.2	Existing analytical models – Literature review on SBS/RSs . . . . .	40
5.1	Notation used in the system description . . . . .	80
5.2	Notation of the random variables used to describe the distributions of the queueing model . . . . .	85
5.3	Notation of the random variables used to describe the distributions of the performance measures . . . . .	99
5.4	Notation used in the system description . . . . .	111
5.5	Notation of the random variables used to describe the distributions of the queueing model . . . . .	115
5.6	Notation of the random variables used to describe the distributions of the performance measures . . . . .	124
6.1	Tested parameter configurations of the tier-captive and tier-to-tier SBS/RSs . . . . .	136
6.2	Overview of the average absolute deviations $ \Delta^{rel} _{avg}$ of $E[T_R]$ , $T_{R,0.95}$ , $E[D_{out}]$ , and $D_{out,0.95}$ . . . . .	143
6.3	Overview of the average absolute deviations $ \Delta^{abs} _{avg}$ of $E[Q_{liftin}]$ , $Q_{liftin,0.95}$ , $E[Q_{ps,1}]$ , and $Q_{ps,1,0.95}$ . . . . .	143
6.4	Overview of the average absolute deviations $ \Delta^{rel} _{avg}$ of $E[T_R]$ , $T_{R,0.95}$ , $E[D_{out}]$ , and $D_{out,0.95}$ . . . . .	150
6.5	Overview of the average absolute deviations $ \Delta^{abs} _{avg}$ of $E[Q_{aisle,S}]$ , $Q_{aisle,S,0.95}$ , $E[Q_{ps,1}]$ , and $Q_{ps,1,0.95}$ . . . . .	151
6.6	Requirements for the tier-captive SBS/RS . . . . .	152
6.7	Application example single-level shuttle . . . . .	156
6.8	Application example multi-level shuttle with two levels . . . . .	159
6.9	Application example multi-level shuttle with three levels . . . . .	160
6.10	Application example multi-level shuttle with four levels . . . . .	160

6.11	Varied parameters of the tier-captive and tier-to-tier SBS/RSSs . . . . .	162
6.12	Possible combinations of the tier-captive and tier-to-tier SBS/RSSs . . . . .	163
6.13	Impact of the variability reduction of random variables and the absence of re-entrant bins on the 95% quantile of the retrieval transaction time . . . . .	166
A.1	Discrete distributions of the inter-arrival times $A_R$ of the tier-captive configurations with different expected values $E[A_R]$ . . . . .	211
A.2	Discrete distributions of the inter-arrival times $A_R$ of the tier-to-tier configurations with different expected values $E[A_R]$ . . . . .	212
A.3	Conveyor distances of the tier-captive and tier-to-tier SBS/RSSs . . . . .	214
B.1	Impact of re-entrance and random variables on the 95% quantile of the retrieval transaction time $T_{R,0.95}$ in the application example – values of $T_{R,0.95}$ per configuration and combination given in seconds . . . . .	215
C.1	Requirements of the tier-to-tier SBS/RSS . . . . .	219
C.2	Application example single-level shuttle . . . . .	220
C.3	Application example multi-level shuttle with two levels . . . . .	221
C.4	Application example multi-level shuttle with three levels . . . . .	222
C.5	Application example multi-level shuttle with four levels . . . . .	222
C.6	Impact of re-entrance and random variables on the 95% quantile of the retrieval transaction time $T_{R,0.95}$ in the application example – values of $T_{R,0.95}$ per configuration and combination given in seconds . . . . .	223

# A Discrete distributions and conveyor data of the validation and numerical evaluation

## A.1 Discrete distributions

Table A.1: Discrete distributions of the inter-arrival times  $A_R$  of the tier-captive configurations with different expected values  $E[A_R]$

$E[A_R]$ [s]	4.0	5.0	6.0	7.0	8.0	9.0
$cv^2[A_R]$	0.25	0.25	0.25	0.25	0.25	0.25
i [s]	$\alpha_{R,i}$	$\alpha_{R,i}$	$\alpha_{R,i}$	$\alpha_{R,i}$	$\alpha_{R,i}$	$\alpha_{R,i}$
0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	0.05621	0.06430	0.03916	0.15881	0.03078	0.05861
2	0.07041	0.03518	0.01171	0.00085	0.06783	0.00000
3	0.44736	0.41034	0.19650	0.00000	0.07421	0.00000
4	0.17123	0.00241	0.10080	0.07136	0.03238	0.01871
5	0.08745	0.00221	0.19431	0.00617	0.08257	0.13280
6	0.00000	0.00563	0.10082	0.07792	0.09159	0.04928
7	0.00000	0.36989	0.06321	0.33539	0.08512	0.17632
8	0.16734	0.00629	0.09995	0.08488	0.09434	0.07723
9	0.00000	0.09719	0.00000	0.00280	0.08188	0.16546
10	0.00000	0.00656	0.09815	0.09218	0.09649	0.08079
11	0.00000	0.00000	0.00000	0.02272	0.07279	0.06645
12	0.00000	0.00000	0.09539	0.10047	0.06608	0.00000
13	0.00000	0.00000	0.00000	0.02395	0.00346	0.00539

A Discrete distributions and conveyor data

---

14	0.00000	0.00000	0.00000	0.02250	0.00000	0.00000
15	0.00000	0.00000	0.00000	0.00000	0.09967	0.00000
16	0.00000	0.00000	0.00000	0.00000	0.02081	0.00000
17	0.00000	0.00000	0.00000	0.00000	0.00000	0.08456
18	0.00000	0.00000	0.00000	0.00000	0.00000	0.08440

Table A.2: Discrete distributions of the inter-arrival times  $A_R$  of the tier-to-tier configurations with different expected values  $E[A_R]$

$E[A_R]$ [s]	15	20	25	30	
	$cv^2[A_R]$	0.25	0.25	0.25	0.25
i [s]	$\alpha_{R,i}$	$\alpha_{R,i}$	$\alpha_{R,i}$	$\alpha_{R,i}$	
0	0.00000	0.00000	0.00000	0.00000	
1	0.05986	0.00050	0.00209	0.00059	
2	0.06418	0.00000	0.00099	0.07137	
3	0.00002	0.14885	0.00029	0.00000	
4	0.00000	0.04849	0.14338	0.05643	
5	0.05434	0.00000	0.00000	0.00189	
6	0.00000	0.00049	0.00032	0.00374	
7	0.12520	0.00099	0.00099	0.00000	
8	0.00000	0.00000	0.00212	0.00000	
9	0.00537	0.00000	0.00358	0.00000	
10	0.00000	0.10525	0.00537	0.00000	
11	0.00000	0.00000	0.00000	0.00000	
12	0.00006	0.00000	0.00000	0.03717	
13	0.00000	0.00010	0.00000	0.00000	
14	0.00005	0.00000	0.00000	0.00000	
15	0.04044	0.00000	0.00000	0.00000	
16	0.00537	0.00000	0.00535	0.00000	
17	0.08545	0.00000	0.00000	0.00000	
18	0.06761	0.00000	0.00000	0.02675	
19	0.17451	0.00000	0.00000	0.00000	
20	0.15017	0.00538	0.00000	0.00000	
21	0.02199	0.00000	0.18179	0.00000	
22	0.02881	0.00000	0.03318	0.00189	
23	0.03297	0.00000	0.03875	0.00936	

---

24	0.03431	0.16967	0.05996	0.01722
25	0.03291	0.14499	0.07429	0.02348
26	0.00974	0.18923	0.08178	0.02815
27	0.00664	0.02997	0.00100	0.02652
28	0.00000	0.04000	0.07604	0.05680
29	0.00000	0.04681	0.06288	0.00925
30	0.00000	0.01044	0.04285	0.11102
31	0.00000	0.03698	0.05310	0.11140
32	0.00000	0.00219	0.00000	0.10381
33	0.00000	0.00000	0.00000	0.04907
34	0.00000	0.00537	0.00000	0.00000
35	0.00000	0.00000	0.00099	0.00056
36	0.00000	0.00000	0.00000	0.05755
37	0.00000	0.00000	0.00000	0.00105
38	0.00000	0.00000	0.00000	0.00123
39	0.00000	0.00000	0.00000	0.00141
40	0.00000	0.01430	0.00000	0.00157
41	0.00000	0.00000	0.00000	0.00170
42	0.00000	0.00000	0.00000	0.00181
43	0.00000	0.00000	0.00099	0.00190
44	0.00000	0.00000	0.00003	0.00197
45	0.00000	0.00000	0.00000	0.00202
46	0.00000	0.00000	0.00536	0.00204
47	0.00000	0.00000	0.00000	0.00204
48	0.00000	0.00000	0.00000	0.00202
49	0.00000	0.00000	0.00000	0.00197
50	0.00000	0.00000	0.12253	0.10323
51	0.00000	0.00000	0.00000	0.00181
52	0.00000	0.00000	0.00000	0.00170
53	0.00000	0.00000	0.00000	0.00157
54	0.00000	0.00000	0.00000	0.00141
55	0.00000	0.00000	0.00000	0.00123
56	0.00000	0.00000	0.00000	0.00103
57	0.00000	0.00000	0.00000	0.00080
58	0.00000	0.00000	0.00000	0.00056
59	0.00000	0.00000	0.00000	0.00000
60	0.00000	0.00000	0.00000	0.05991

---

## A.2 Distances of the conveyor system used in the simulation

Table A.3: Conveyor distances of the tier-captive and tier-to-tier SBS/RSs

Part of the conveyor	Length
Distance from the merge of replenishment and re-entrant storage transactions to the split into the first aisle	$2m$
Distance between two consecutive splits of the storage transactions into the aisles	$2m$
Distance from the split into the aisles to the (incoming) lifts	$3.5m$
Distance from the (outgoing) lifts to the merge with the other retrieval transactions on the conveyor	$3.5m$
Distance between two consecutive merges of the retrieval transactions coming from the aisles	$2m$
Distance from the last merge of the retrieval transactions to the split into the first picking station	$8m$
Distance from the split into the picking station to the picking station	$3.5m$
Distance from the picking station to the merge with the other transactions on the conveyor	$3.5m$
Distance for transactions bypassing a picking station	$2m$
Distance between two consecutive picking stations	$2m$
Distance between the last picking station and the split into transactions leaving and re-entering the system	$8m$
Distance between the split of transactions re-entering the system and the merge with the replenishment transactions	$2m$

## B Application example of a tier-captive system: additional data

Table B.1: Impact of re-entrance and random variables on the 95% quantile of the retrieval transaction time  $T_{R,0.95}$  in the application example – values of  $T_{R,0.95}$  per configuration and combination given in seconds

Config.	Combination							
	1	2	3	4	5	6	7	8
1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4	220	219	219	218	220	220	219	219
5	191	190	190	189	191	191	190	190
6	170	170	169	169	170	170	169	169
7	155	154	154	154	155	155	154	154
8	143	143	142	142	143	143	142	142
9	134	134	133	133	134	134	134	134
10	127	127	126	126	127	127	126	126
11	122	121	121	121	122	122	121	121
12	118	118	117	117	118	118	117	117
13	114	114	114	114	114	114	114	114
14	111	111	111	111	111	111	111	111
15	204	203	203	202	204	204	203	203
16	170	170	169	169	170	170	169	169
17	148	148	147	147	148	148	147	147
18	133	132	132	132	133	133	132	132
19	121	121	120	120	121	121	120	120

B Application example of a tier-captive system

---

20	111	110	110	110	111	111	110	110
21	103	103	103	103	103	103	103	103
22	97	97	96	96	97	97	96	96
23	92	92	92	92	92	92	92	92
24	88	87	87	87	88	88	87	87
25	84	84	84	84	84	84	84	84
26	82	82	81	81	82	82	81	81
27	79	79	79	78	79	79	79	79
28	77	77	77	77	77	77	77	77
29	75	75	75	75	75	75	75	75
30	225	224	224	223	225	225	224	224
31	175	175	174	174	175	175	175	174
32	146	146	146	145	147	146	146	146
33	127	127	126	126	127	127	126	126
34	113	113	112	112	113	113	112	112
35	103	102	102	102	103	103	102	102
36	94	94	93	93	94	94	93	93
37	87	87	87	87	87	87	87	87
38	82	82	82	82	82	82	82	82
39	78	78	78	78	78	78	78	78
40	74	74	74	74	74	74	74	74
41	71	71	71	71	71	71	71	71
42	68	68	68	68	68	68	68	68
43	66	66	66	66	66	66	66	66
44	65	65	65	65	65	65	65	65
45	63	63	63	63	63	63	63	63
46	61	61	61	61	61	61	61	61
47	61	61	60	60	61	61	60	60
48	n.a.							
49	n.a.							
50	423	422	419	418	423	423	420	420
51	247	246	245	244	247	247	245	245
52	183	183	182	181	183	183	182	182
53	153	153	152	152	153	153	152	152
54	n.a.							
55	281	280	278	277	281	281	279	278
56	176	176	175	174	176	176	175	175
57	135	134	134	133	135	135	134	134

58	113	112	112	112	113	113	112	112
59	99	99	99	98	99	99	99	99
60	90	90	89	89	90	90	89	89
61	n.a.							
62	420	418	416	414	421	420	417	417
63	191	190	190	189	192	191	190	190
64	133	132	132	131	133	133	132	132
65	107	107	106	106	107	107	106	106
66	90	90	90	89	90	90	90	90
67	80	80	79	79	80	80	79	79
68	73	73	73	73	73	73	73	73
69	68	67	67	67	68	68	67	67
70	n.a.							
71	n.a.							
72	n.a.							
73	542	541	536	535	542	542	537	537
74	232	231	230	229	232	232	230	230
75	n.a.							
76	540	538	535	532	541	541	536	535
77	194	194	192	192	195	194	193	192
78	129	129	128	128	129	129	128	128
79	102	102	101	101	102	102	101	101
80	n.a.							
81	364	362	360	358	365	364	361	360
82	152	151	150	150	152	152	151	150
83	105	105	104	104	105	105	104	104
84	84	84	84	83	84	84	84	84
85	73	73	73	72	73	73	73	73
86	n.a.							
87	n.a.							
88	n.a.							
89	n.a.							
90	196	196	194	193	197	197	195	194
91	n.a.							
92	n.a.							
93	162	161	160	160	162	162	161	160
94	100	100	99	99	100	100	99	99



# C Application example of a tier-to-tier system: additional data

Table C.1: Requirements of the tier-to-tier SBS/RS

Parameter	Values
Storage capacity	$Cap = 20,000$
Maximum height the system	$max_H = 10m$
Maximum length of the system	$max_L = 100m$
Maximum width of the system	$max_W = 10m$
Throughput of the retrieval transactions	$\lambda_R = 100 \frac{1}{h}$
Maximum utilization	$\rho_{max} = 0.9$
Maximum value of the 95% quantile of the retrieval transaction time	$T_{R,0.95} = 120s$

Table C.2: Application example single-level shuttle

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [m <sup>2</sup> ]	$n_{L,tot}$	$n_V$	$\rho_a$	$T_{R,0.95}$ [s]
1	2, 1, 25, 200	20,000	400.00	2	2	1.71	n.a.
2	2, 1, 26, 193	20,072	386.00	2	2	1.67	n.a.
3	2, 1, 27, 186	20,088	372.00	2	2	1.62	n.a.
4	3, 1, 17, 197	20,094	591.00	3	3	1.12	n.a.
5	3, 1, 18, 186	20,088	558.00	3	3	1.07	n.a.
6	3, 1, 19, 176	20,064	528.00	3	3	1.02	n.a.
7	3, 1, 20, 167	20,040	501.00	3	3	0.98	n.a.
8	3, 1, 21, 159	20,034	477.00	3	3	0.95	n.a.
9	3, 1, 22, 152	20,064	456.00	3	3	0.92	n.a.
10	3, 1, 23, 145	20,010	435.00	3	3	0.89	754
11	3, 1, 24, 139	20,016	417.00	3	3	0.86	596
12	3, 1, 25, 134	20,100	402.00	3	3	0.84	503
13	3, 1, 26, 129	20,124	387.00	3	3	0.82	433
14	3, 1, 27, 124	20,088	372.00	3	3	0.80	377
15	4, 1, 13, 193	20,072	772.00	4	4	0.82	598
16	4, 1, 14, 179	20,048	716.00	4	4	0.77	452
17	4, 1, 15, 167	20,040	668.00	4	4	0.73	366
18	4, 1, 16, 157	20,096	628.00	4	4	0.70	312
19	4, 1, 17, 148	20,128	592.00	4	4	0.67	272
20	4, 1, 18, 139	20,016	556.00	4	4	0.64	239
21	4, 1, 19, 132	20,064	528.00	4	4	0.62	216
22	4, 1, 20, 125	20,000	500.00	4	4	0.59	196
23	4, 1, 21, 120	20,160	480.00	4	4	0.58	183
24	4, 1, 22, 114	20,064	456.00	4	4	0.56	169
25	4, 1, 23, 109	20,056	436.00	4	4	0.54	158
26	4, 1, 24, 105	20,160	420.00	4	4	0.53	150
27	4, 1, 25, 100	20,000	400.00	4	4	0.51	140
28	4, 1, 26, 97	20,176	388.00	4	4	0.50	135
29	4, 1, 27, 93	20,088	372.00	4	4	0.49	128
30	5, 1, 10, 200	20,000	1000.00	5	5	0.67	348
31	5, 1, 11, 182	20,020	910.00	5	5	0.62	284
32	5, 1, 12, 167	20,040	835.00	5	5	0.58	240
33	5, 1, 13, 154	20,020	770.00	5	5	0.55	209
34	5, 1, 14, 143	20,020	715.00	5	5	0.52	186
35	5, 1, 15, 134	20,100	670.00	5	5	0.49	169
36	5, 1, 16, 125	20,000	625.00	5	5	0.47	153
37	5, 1, 17, 118	20,060	590.00	5	5	0.45	142
38	5, 1, 18, 112	20,160	560.00	5	5	0.44	133
39	5, 1, 19, 106	20,140	530.00	5	5	0.42	124
40	<b>5, 1, 20, 100</b>	<b>20,000</b>	<b>500.00</b>	<b>5</b>	<b>5</b>	<b>0.41</b>	<b>116</b>

<b>41</b>	<b>5, 1, 21, 96</b>	<b>20,160</b>	<b>480.00</b>	<b>5</b>	<b>5</b>	<b>0.40</b>	<b>111</b>
<b>42</b>	<b>5, 1, 22, 91</b>	<b>20,020</b>	<b>455.00</b>	<b>5</b>	<b>5</b>	<b>0.38</b>	<b>105</b>
<b>43</b>	<b>5, 1, 23, 87</b>	<b>20,010</b>	<b>435.00</b>	<b>5</b>	<b>5</b>	<b>0.37</b>	<b>100</b>
<b>44</b>	<b>5, 1, 24, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>5</b>	<b>5</b>	<b>0.36</b>	<b>96</b>
<b>45</b>	<b>5, 1, 25, 80</b>	<b>20,000</b>	<b>400.00</b>	<b>5</b>	<b>5</b>	<b>0.35</b>	<b>92</b>
<b>46</b>	<b>5, 1, 26, 77</b>	<b>20,020</b>	<b>385.00</b>	<b>5</b>	<b>5</b>	<b>0.35</b>	<b>89</b>
<b>47</b>	<b>5, 1, 27, 75</b>	<b>20,250</b>	<b>375.00</b>	<b>5</b>	<b>5</b>	<b>0.34</b>	<b>87</b>

Table C.3: Application example multi-level shuttle with two levels

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [ $m^2$ ]	$n_{L,tot}$	$n_V$	$\rho_a$	$T_{R,0.95}$ [s]
48	2, 2, 13, 193	20,072	386.00	2	2	1.66	n.a.
49	3, 2, 9, 186	20,088	558.00	3	3	1.06	n.a.
50	3, 2, 10, 167	20,040	501.00	3	3	0.98	n.a.
51	3, 2, 11, 152	20,064	456.00	3	3	0.91	n.a.
52	3, 2, 12, 139	20,016	417.00	3	3	0.85	568
53	3, 2, 13, 129	20,124	387.00	3	3	0.81	418
54	4, 2, 7, 179	20,048	716.00	4	4	0.76	431
55	4, 2, 8, 157	20,096	628.00	4	4	0.69	302
56	4, 2, 9, 139	20,016	556.00	4	4	0.63	233
57	4, 2, 10, 125	20,000	500.00	4	4	0.59	192
58	4, 2, 11, 114	20,064	456.00	4	4	0.55	166
59	4, 2, 12, 105	20,160	420.00	4	4	0.52	148
60	4, 2, 13, 97	20,176	388.00	4	4	0.50	133
61	5, 2, 5, 200	20,000	1000.00	5	5	0.66	332
62	5, 2, 6, 167	20,040	835.00	5	5	0.57	233
63	5, 2, 7, 143	20,020	715.00	5	5	0.51	181
64	5, 2, 8, 125	20,000	625.00	5	5	0.47	150
65	5, 2, 9, 112	20,160	560.00	5	5	0.43	131
<b>66</b>	<b>5, 2, 10, 100</b>	<b>20,000</b>	<b>500.00</b>	<b>5</b>	<b>5</b>	<b>0.40</b>	<b>115</b>
<b>67</b>	<b>5, 2, 11, 91</b>	<b>20,020</b>	<b>455.00</b>	<b>5</b>	<b>5</b>	<b>0.38</b>	<b>103</b>
<b>68</b>	<b>5, 2, 12, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>5</b>	<b>5</b>	<b>0.36</b>	<b>95</b>
<b>69</b>	<b>5, 2, 13, 77</b>	<b>20,020</b>	<b>385.00</b>	<b>5</b>	<b>5</b>	<b>0.34</b>	<b>88</b>

C Application example of a tier-to-tier system

---

Table C.4: Application example multi-level shuttle with three levels

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [m <sup>2</sup> ]	$n_{L,tot}$	$n_V$	$\rho_a$	$T_{R,0.95}$ [s]
70	2, 3, 9, 186	20,088	372.00	2	2	1.60	n.a.
71	3, 3, 6, 186	20,088	558.00	3	3	1.05	n.a.
72	3, 3, 7, 159	20,034	477.00	3	3	0.93	n.a.
73	3, 3, 8, 139	20,016	417.00	3	3	0.85	544
74	3, 3, 9, 124	20,088	372.00	3	3	0.79	356
75	4, 3, 5, 167	20,040	668.00	4	4	0.72	340
76	4, 3, 6, 139	20,016	556.00	4	4	0.63	228
77	4, 3, 7, 120	20,160	480.00	4	4	0.57	177
78	4, 3, 8, 105	20,160	420.00	4	4	0.52	146
79	4, 3, 9, 93	20,088	372.00	4	4	0.48	125
80	5, 3, 4, 167	20,040	835.00	5	5	0.57	226
81	5, 3, 5, 134	20,100	670.00	5	5	0.48	162
82	5, 3, 6, 112	20,160	560.00	5	5	0.43	129
83	<b>5, 3, 7, 96</b>	<b>20,160</b>	<b>480.00</b>	<b>5</b>	<b>5</b>	<b>0.39</b>	<b>108</b>
84	<b>5, 3, 8, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>5</b>	<b>5</b>	<b>0.36</b>	<b>94</b>
85	<b>5, 3, 9, 75</b>	<b>20,250</b>	<b>375.00</b>	<b>5</b>	<b>5</b>	<b>0.34</b>	<b>85</b>

Table C.5: Application example multi-level shuttle with four levels

No.	$n_a, n_l, n_t, n_c$	Cap	Footprint [m <sup>2</sup> ]	$n_{L,tot}$	$n_V$	$\rho_a$	$T_{R,0.95}$ [s]
86	3, 4, 5, 167	20,040	501.00	3	3	0.96	n.a.
87	3, 4, 6, 139	20,016	417.00	3	3	0.84	521
88	4, 4, 4, 157	20,096	628.00	4	4	0.68	283
89	4, 4, 5, 125	20,000	500.00	4	4	0.58	185
90	4, 4, 6, 105	20,160	420.00	4	4	0.52	144
91	5, 4, 3, 167	20,040	835.00	5	5	0.56	218
92	5, 4, 4, 125	20,000	625.00	5	5	0.46	144
93	<b>5, 4, 5, 100</b>	<b>20,000</b>	<b>500.00</b>	<b>5</b>	<b>5</b>	<b>0.40</b>	<b>111</b>
94	<b>5, 4, 6, 84</b>	<b>20,160</b>	<b>420.00</b>	<b>5</b>	<b>5</b>	<b>0.36</b>	<b>93</b>

Table C.6: Impact of re-entrance and random variables on the 95% quantile of the retrieval transaction time  $T_{R,0.95}$  in the application example – values of  $T_{R,0.95}$  per configuration and combination given in seconds

Config.	Combination							
	1	2	3	4	5	6	7	8
1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
5	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
7	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
8	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
9	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
10	790	689	689	605	771	761	673	664
11	624	544	544	478	609	601	532	525
12	527	459	459	404	515	507	450	443
13	452	395	395	347	442	436	387	381
14	394	344	344	302	385	380	337	332
15	618	561	561	510	610	604	553	548
16	467	424	424	386	461	456	418	414
17	378	343	343	313	373	370	339	336
18	322	292	292	266	318	315	289	286
19	281	255	255	232	278	275	252	250
20	246	224	224	204	244	241	221	219
21	223	203	203	184	221	218	200	198
22	202	184	184	167	200	198	182	180
23	189	172	172	156	187	185	170	168
24	174	158	158	144	173	170	157	155
25	163	148	148	135	161	159	147	145
26	154	140	140	128	153	151	139	137
27	144	131	131	120	143	141	130	128
28	139	126	126	115	138	136	125	123
29	131	120	120	109	130	129	119	117
30	357	331	331	308	355	352	329	326
31	291	270	270	251	289	286	268	266
32	246	229	229	212	245	243	227	225
33	214	199	199	185	213	211	197	196

C Application example of a tier-to-tier system

---

34	190	177	177	164	189	187	175	174
35	173	160	160	149	172	170	159	158
36	157	145	145	135	156	154	145	143
37	145	135	135	125	144	143	134	133
38	136	126	126	117	135	134	125	124
39	127	118	118	109	126	125	117	116
40	119	110	110	102	118	117	110	108
41	113	105	105	97	113	112	105	104
42	107	99	99	92	106	105	99	98
43	102	95	95	88	102	100	94	93
44	99	91	91	84	98	97	91	90
45	94	87	87	80	93	92	87	85
46	90	84	84	77	90	89	83	82
47	88	82	82	75	88	87	82	80
48	n.a.							
49	n.a.							
50	n.a.							
51	n.a.							
52	594	519	519	455	581	573	507	500
53	437	382	382	335	428	421	374	368
54	445	404	404	368	439	435	399	395
55	312	283	283	258	308	305	280	277
56	240	218	218	199	238	235	216	214
57	198	180	180	164	196	194	178	176
58	171	156	156	142	170	168	154	152
59	152	138	138	126	151	149	137	135
60	137	125	125	114	136	134	123	122
61	341	316	316	293	338	335	313	311
62	239	221	221	206	237	235	220	218
63	186	172	172	160	185	183	171	170
64	154	143	143	132	153	151	142	140
65	134	124	124	115	133	132	123	122
66	117	109	109	101	116	115	108	107
67	106	98	98	91	105	104	98	96
68	97	90	90	83	97	96	90	89
69	90	83	83	77	89	88	83	81
70	n.a.							
71	n.a.							

72	n.a.							
73	569	497	497	436	557	549	486	479
74	372	325	325	286	364	359	318	314
75	351	319	319	290	347	343	315	312
76	235	213	213	194	232	230	211	209
77	183	166	166	151	181	179	164	162
78	150	137	137	124	149	147	135	134
79	128	117	117	107	127	126	116	114
80	232	215	215	200	230	228	213	212
81	166	154	154	143	165	163	153	151
82	132	122	122	113	131	130	122	120
83	111	103	103	95	110	109	102	101
84	96	89	89	83	96	95	89	88
85	87	80	80	74	86	85	80	79
86	n.a.							
87	545	476	476	418	533	525	466	459
88	292	266	266	242	289	286	263	260
89	191	173	173	158	189	187	172	170
90	148	135	135	123	147	145	133	132
91	224	208	208	193	222	220	206	204
92	148	137	137	127	147	146	136	135
93	114	106	106	98	113	112	105	104
94	95	88	88	82	95	94	88	87

**Wissenschaftliche Berichte des Instituts für Fördertechnik und  
Logistiksysteme des Karlsruher Instituts für Technologie (KIT)  
Prof. Dr.-Ing. Kai Furmans [Hrsg.]**



Shuttle-based storage and retrieval systems (SBS/RSs) are an important part of today's warehouses. Due to their ability to generate a high throughput, they represent an interesting alternative to traditional automated storage and retrieval systems. The fast technological development in the field of SBS/RSs creates the need for performance evaluation tools that can be used during the decision making process.

Hence, a new approach is developed that can be applied to model different configurations of SBS/RSs. The approach is based on the modeling of SBS/RSs as discrete-time open queueing networks and yields the complete probability distributions of the performance measures. Given the system requirements such as storage capacity, throughput, and service level, it is shown in this work how to determine a suitable system design that fulfills the requirements. Moreover, the impact of multi-level shuttles, re-entrant bins, and the variability of the random variables is analyzed.

ISSN 0171-2772

ISBN 978-3-7315-0742-0

ISBN 978-3-7315-0742-0



9 783731 507420 >