3. RISK ANALYSIS OF GENERALIZED JACKSON NETWORKS

This chapter is based on [171, 31].

Jackson networks, introduced by J. R. Jackson in [107, 108], are robust models for analysing complex networks of nodes, where each node represents a server with a queue. Customers in the network follow a fixed routing matrix and the key characterization of Jackson networks is that the equilibrium distribution of the queues length vector can be easily computed via a product-form expression.

In this chapter we study generalized Jackson networks with single server stations, where nodes may have an infinite supply of work. We allow in our model simultaneous breakdown of groups of servers, and group repair strategies, thereby capturing the feature that repairing several servers simultaneously may lead to more efficient repair actions and thus may reduce the repair time. We will establish the existence of a steady-state distribution of the queue length vector at stable nodes for different type of failure regimes. As it turns out, the distribution of the failure/repair regime and the steady-state distribution of the queue length vector at stable nodes decouple in a product-form way, and we will provide closed-form solutions for the classical performance measures such as the long-run throughput or mean waiting time at a station.

With these closed-form solutions at hand, we show how parameter insecurity can be incorporated into the model as well. The motivation for this extension stems from reliability where typically the distribution of the time between breakdowns is hard to estimate and therefore the parameters characterizing the distribution are typically not (exactly) known. We illustrate the impact of our results with an analysis of parameter insecurity of the throughput of the systems for different breakdown regimes. Specifically, we will provide an approach for evaluating the value at risk under parameter insecurity.

The chapter is organized as follows. Section 3.1 provides an introduction to generalized Jackson networks and our risk analysis research. In Section 3.2 we present a motivating example, and we will introduce the basic notation and concepts with the help of this example. At the end of Section 3.2 a summary of the introduced notation can be found. A literature review is provided in Section 3.3. The main technical analysis is carried out in Section 3.4. In Section 3.4.4, we provide explicit solutions for availability and performance measures. Eventually, Section 3.5 is devoted to robustness analysis of the performance measures.

3.1 Introduction

Design and analysis of stochastic networks is often challenged by the fact that the exact specifications of the network are either not known, or by lack of sufficient data for
calibrating the model. In this chapter we assume that the physical layout of the network, i.e., the number of nodes, and the topology are known, as well as mean service times, mean inter-arrival times of customers at the network and routing decisions. Provided that this rough information constitute the only available data in advance, arguments exploiting entropy properties lead to use so-called product-form models as conservative first order models. Indeed, if mean service times and mean inter-arrival times are given, the exponential distribution is known to maximize the entropy over all distributions with support $\mathbb{R}_+:=[0,\infty)$ and these means\(^1\), see, for example, [160, 138]. Furthermore, with obtained mean local flow and service rate it is known that a product-form solution maximizes the entropy of the stationary queue length distribution [68, 186]. Therefore, product-form solutions are conservative and robust models with respect to model insecurity in the service time and inter-arrival time distributions. Hence, working with a model with exponentially distributed service times and inter-arrival times that does have a product-form solution for the stationary queue length distribution, provides a robust model for performance analysis. This insight motivates our research into Jackson networks (henceforth JN), which are a well established class of models in, e.g., production, telecommunication, computer systems; for surveys see [113] and [54]. JN have the desirable property that the distribution of the stationary queue length vector is of product-form, which allows for quick numerical evaluation of performance measures, such as the mean queue length, mean sojourn times and the throughput at nodes. Unfortunately, while having the desirable robustness property the class of JN is rather limited and for this reasons we extend in this chapter the classical JN framework in two directions, namely to networks with infinite supply nodes and breakdowns:

- We allow for servers to have infinite supply. Infinite supply has the aim to utilize the capacity of a server to the fullest. For example, in service centre models it is typically assumed that an agent, when not answering a call, switches to low priority works such as answering email and administrative duties. In a traffic model of a highway network, an infinite supply node represents a highway segment with an on-ramp during a rush-hour period where a constant flow of vehicles requiring access to the highway is present. In production processes, an additional inventory with raw material guarantees that the machine will not be idle even when there is sometimes no external demand.

- We allow for breakdown and repair of individual servers and for simultaneous breakdown and repair of groups of servers (i.e., repair can be grouped). This allows (i) to model simultaneous breakdown of groups of servers, and (ii) model group repair strategies. For example, repairing several servers simultaneously may lead to more efficient repair actions and thus may reduce the repair time.

- We allow for unstable nodes in the network, where instability of nodes may be due to overload generated by nodes without infinite supply or nodes with infinite supply or both.

---

\(^1\) Loosely speaking, entropy can be seen in this context as the amount of uncertainty. Results based on distributions given certain means with maximal entropy should be ‘least surprising’ in terms of predictions that follow from the model. Therefore the most conservative probabilistic model for service times (or inter-arrival times) with support $\mathbb{R}_+$ consistent with a given mean value is an exponential distribution.
While the first and the third feature, and the combination thereof were already dealt with in the literature (more details later on in Subsection 3.3), the interaction of Jackson networks with infinite supply, instability of nodes, and reliability issues has to the best of our knowledge not found the deserved investigations.

To fill this gap in the literature we will show in the first part of this chapter that JN with the above additional feature still have product-form property and that the classical performance measures, such as, throughput, mean queue length, and mean service time, are analytical in the the system parameters.

For the second part of the chapter, we take as starting point the simple closed-form solutions for the standard performance measures of the extended JN. Let, for example, \( L(\theta) \) be the throughput of the network and let \( \theta \) denote the mean service time at some server of the network. Suppose that \( \theta \) is not exactly known but data are available for estimating \( \theta \). Hence, based on the data, we have an empirical distribution available for the true value of \( \theta \).

Evaluating the performance for a fixed value of \( \theta \) does not reflect the fact that there is parameter insecurity with respect to \( \theta \) and we will therefore assume \( \theta \) to be a random variable the distribution of which is obtained from the sample statistic. To obtain a responsible performance evaluation, we will consider the value at risk of \( L(\theta) \) incurred by the distribution of \( \theta \). In the second part of the chapter, we will provide a power series approximation for the value at risk of \( L(\theta) \) the elements of which are easily computable. This value at risk approach is of particularly interest for models with breakdowns. Indeed, data on the lifetime of a component that is prone to breakdowns is only available in censored form and estimating mean lifetimes of components is a statistical challenge. Therefore, one is typically confronted with uncertainty about the true value of the parameters defining the distributions of the time between breakdowns, in our case the failure rate. This is known as parameter uncertainty in the literature, see, for example, [88], for a discussion on integration of parameter uncertainty into queueing models and [94] for a discussion on parameter insecurity from a broader perspective.

### 3.2 Motivating Example

We explain the basic setup for our analysis with the following example. Consider the Jackson network with \( J = 7 \) nodes, depicted in Figure 3.1, where we assume throughout that servers are first come first served (FCFS) single-server stations with infinite waiting capacity. The set of all nodes is denoted by \( \tilde{J} = \{1, \ldots, J\} \). Throughout this chapter the cardinality of a set \( A \) will be denoted with \( |A| \), so \( |	ilde{J}| = J \). In addition we will denote the outside world, that is, the source as well as the sink, as node 0. We identify the labelling of the service rate \( \mu_i \) with the server labels, so that \( \mu_i \) refers to server \( i \). There are two arrival streams with arrival rates \( \lambda_1 \) and \( \lambda_2 \), respectively, and throughout this chapter it is assumed that arrival processes are of Poisson type. Routing is Markovian and possible routes are indicated by arrows. Specifically, the probability that a customer after finishing service at server, say, \( i \in \tilde{J} \), moves to node \( j \in \tilde{J} \cup \{0\} \), is given by \( r(i, j) \).

In the same manner, if we denote the total arrival rate of customers from outside the system with \( \lambda = \lambda_1 + \lambda_2 \) then with probability \( r(0, i) = \lambda_i / \lambda \) an arriving customer enters the network at server \( i \). These probabilities are captured in routing matrix

\[
R = \{r(i, j) : i, j \in \tilde{J} \cup \{0\}\} 
\]
Apart from these classical features of a Jackson network, we assume that there is a set of nodes $V$ such that each $j \in V$ has infinite supply. This is indicated in the figure by dotted arrows. We complement this definition by introducing $W = \tilde{J} \setminus V$ as the set of all nodes without infinite supply. In this example, $V = \{1, 5, 7\}$ and $W = \{2, 3, 4, 6\}$. For $j \in V$, customers in the infinite supply chain have low priority, where customers arriving either from the outside or from another server have high priority with pre-emptive-resume regime: service of a low priority customer is interrupted as soon as a high priority customer arrives. When a low priority customer is served and fed into the network, he becomes a high priority customer. Routing decisions, service times, and inter-arrival times are assumed to be mutually independent. The network displayed in Figure 3.1 is inspired by a model of a highway network. Nodes represent road segments and nodes with infinite supply model road segments with an on-ramp, where it is assumed that there is constant flow of incoming traffic to the on-ramp. Note that in queueing models for traffic systems one typically only models the flow in one direction and fits the service rates of the queues to the traffic characteristics, see, for example, [183, 182].

Suppose that $\lambda_1 < \mu_3 < \mu_1$. Then, without infinite supply at node 1, node 3 is stable in the classical definition as the rate with which customers arrive to node 3 is smaller than the service rate. In case of infinite supply at node 1, however, node 1 acts as Poisson source and the incoming traffic rate at node 3, which is then $\mu_1$, is larger than the service rate, which causes node 3 to become unstable. Whether a node is stable or not can be decided from the traffic equations of the network and details are provided in Section 3.4.2. We let $S \subseteq \tilde{J}$ denote the set of stable nodes and by $U$ the
set of unstable nodes.

To complete this introductory example, we denote the set of nodes that are unreliable by $D$. For example, take $D = \{3, 5\}$. Breakdown and repair can follow complicated schemes: nodes may break down isolated or in groups, and repair may happen similarly. It is not required that nodes which are broken down simultaneously are repaired at the same time. We make the simplifying assumption that breakdown and repair intensities do not depend on queue lengths. As will become clear from our analysis, in this model stability of nodes is independent of the breakdown/repair regime. In other words, it is not possible to “create” instability in a network by choosing a misguided repair action.

**Definition 3.1.** Let $\mathcal{P}(D)$ be the power set of $D$. If the nodes in $I \in \mathcal{P}(D)$ are broken down, then

1. if $I \subset H \subseteq D$, the nodes in $H \setminus I$ break down with rate $\alpha(I, H) \geq 0$,

2. if $K \subset I$, the nodes in $I \setminus K$ are repaired with rate $\beta(I, K) \geq 0$.

$\alpha(\cdot, \cdot), \beta(\cdot, \cdot)$ may be constructed from any pair of non-negative functions $A, B : \mathcal{P}(D) \to [0, \infty)$, subject to $A(\emptyset) = B(\emptyset) = 1$ and with $A(H)/A(I) < \infty$, for all $I \subset H \subseteq D$, and $B(I)/B(K) < \infty$, for all $K \subset I$ (where we set $0/0 = 0$).

With these functions we set for all subsets of down nodes $I \subseteq D$

$$
\alpha(I, H) = \frac{A(H)}{A(I)}, \quad I \subset H \subseteq D, \quad \text{and} \quad \beta(I, K) = \frac{B(I)}{B(K)}, \quad K \subset I. \quad (3.1)
$$

The parametric form (3.1) of the breakdown and repair rates is selected because it stems from a versatile recipe to construct correlated multidimensional birth-death processes. A statistical procedure to check whether this form is justified is to determine in a first step all possible values $A(I) = \alpha(\emptyset, I)$ and $B(I) = \beta(I, \emptyset)$, $I \subseteq D$, and then to check stepwise (3.1).

In our example, there are possible breakdown scenarios $\emptyset, \{3\}, \{5\}, \{3, 5\}$. We let $A(\emptyset) = B(\emptyset) = 1$ and assume that the rate with which a breakdown of server $i = 3, 5$ occurs is given by $A(\{i\}) = \tau_i$, and the corresponding repair rate by $B(\{i\}) = \rho_i$. In the same vein, let $A(\{3, 5\}) = \tau_3 + \tau_5$ be the rate with which the operating system enters breakdown state $\{3, 5\}$, and let $B(\{3, 5\}) = 2 \min(\rho_3, \rho_5)$ be the rate with which the systems jumps from breakdown state $\{3, 5\}$ to the state $\emptyset$ with all servers operating. The rates $A(I)$ and $B(I)$ are the basic input data for our model. Following (3.1) we now construct transition rate mappings $\alpha$ and $\beta$ covering all possible intermediate state transitions. More specifically, for $i \in D$ let $\alpha(\emptyset, \{i\}) = \tau_i$ and $\beta(\{i\}, \emptyset) = \rho_i$, $\alpha(\{i\}, \{3, 5\}) = \tau_3 + \tau_5$ and $\beta(\{3, 5\}, \emptyset) = 2 \min(\rho_3, \rho_5)$. Eventually, given that node $i \in D$ is broken down, the breakdown rate of the other node, say, $j$ is given by

$$
\alpha(\{i\}, \{3, 5\}) = \frac{\tau_3 + \tau_5}{\tau_i}
$$

whereas from breakdown scenario $\{3, 5\}$ node $i \in D$ alone has repair rate

$$
\beta(\{3, 5\}, \{j\}) = \frac{2 \min(\rho_3, \rho_5)}{\rho_j}.
$$

All other values for breakdown and repair rates are zero.
Note that these breakdown and repair rates from (3.1) define a generator for a Markov process \( Y = \{ Y_t : t \geq 0 \} \) on state space \( \mathcal{P}(D) \). By inspection we see that

\[
\pi := (\pi(K) := A(K)/B(K), \ K \in \mathcal{P}(D))
\]

fulfills

\[
\pi(K) \cdot \alpha(K, K \cup G) = \pi(K \cup G) \cdot \beta(K \cup G, K)
\]

for all \( K, G \in \mathcal{P}(D) \) which implies that, after normalization, \( \pi \) is the steady-state of the breakdown and repair process. Even more, we have proved that \( Y \) is reversible.

In our example, we have only three possible breakdown scenarios and the repair process has states \( \emptyset, \{3\}, \{5\}, \{3, 5\} \). The stationary distribution \( \pi \) of \( Y \) equals

\[
\pi(\emptyset) = c^{-1}, \ \pi(\{3\}) = c^{-1} \frac{\tau_3}{\rho_3}, \ \pi(\{5\}) = c^{-1} \frac{\tau_5}{\rho_5}, \ \pi(\{3, 5\}) = c^{-1} \frac{\tau_3 + \tau_5}{2 \min(\rho_3, \rho_5)},
\]

with normalizing constant

\[
c = 1 + \frac{\tau_3}{\rho_3} + \frac{\tau_5}{\rho_5} + \frac{\tau_3 + \tau_5}{2 \min(\rho_3, \rho_5)}.
\]

**Summary of Notation**

The following notation is introduced in the preceding motivating example and will be used throughout this chapter:

- \( J \) is the number of nodes in the network.
- \( \tilde{J} = \{1, 2, \ldots, J\} \) is the set of servers in the network (\( |\tilde{J}| = J \), i.e., the cardinality of a set is denoted by \(|\cdot|\)).
- \( \mu_i \) is the service rate at server \( i \in \tilde{J} \).
- \( \lambda_i \) is the arrival rate of customers from outside the network at server \( i \in \tilde{J} \).
- \( \lambda = \sum_{i \in J} \lambda_i \) is the total arrival rate from outside the network.
- \( R = (r(i, j) : i, j \in \tilde{J} \cup \{0\}) \) is the routing matrix (stochastic and irreducible), where \( r(i, j) \) gives the probability that a customer ‘finished’ at node \( i \) moves to node \( j \). Node 0 represents the source and sink of customers.
- \( V \) is the set of servers *with* infinite supply.
- \( W = \tilde{J} \setminus V \) is the set of servers *without* infinite supply.
- \( S \) is the set of *stable* servers.
- \( U = \tilde{J} \setminus S \) is the set of *unstable* servers.
- \( D \) is the set of unreliable servers that can breakdown.
- \( \mathcal{P}(D) \) is the power set of \( D \).
- \( A(\cdot) \) is a function corresponding to breakdown rates \( \alpha(\cdot, \cdot) \).
- \( B(\cdot) \) is a function corresponding to repair rates \( \beta(\cdot, \cdot) \).
3.3 Literature Review and Related Work

Investigation of generalized Jackson networks with infinite supply has recently found much interest in the literature and it turned out that the feature of infinite supply makes analysis of the network considerably harder than that of classical product-form networks of the BCMP, [26], and Kelly type, [112].

Infinite supply of lower-priority work (also known as infinite virtual queues or in short IVQ) is used frequently, e.g., in [136] in a $M/G/1$ queueing system to utilize idle times. Recent works using this concept of infinite supply are, e.g., [81] where generalized Jackson networks are considered or [126] where a push-pull network with infinite supply is investigated.

A special class of multi-class queueing networks with infinite virtual queues has been introduced in [127] and [4]. For single-class ergodic networks of Jackson type with infinite supply of work at some nodes G. Weiss [188] has obtained a product-form solution of the steady-state queue length distribution at nodes without infinite supply. He discussed as an example a particular computer communication system that works according to MAN (metropolitan area network) Ethernet RPR (resilient packet ring), where ring traffic has priority over the traffic generated at nodes.

Another application from a different field where such a model fits is in wireless sensor networks. The nodes (sensors) continuously sense their environment and have to forward the data to a central station (sink). This is usually not possible by direct communication, so the nodes act additionally as transmission stations for data from other sensor nodes. If forwarding transmissions from other nodes has priority, its own data constitute the infinite buffer which generates the infinite supply for the node.

The work of Guo [81] and of Guo et al. [82] is on general multi-class queueing networks with IVQ under different scheduling policies for the servers. These policies guide the nodes’ decisions how to dedicate their activities to either the regular standard queues or the infinite virtual queues. The key research question is the interplay of the production of jobs from the IVQ and stability of the standard queues.

Another class of models where additional work is added whenever a server becomes idle are queues with vacations. If a server observes an empty queue “he goes away to serve at some other place a customer”, and returns thereafter. If he finds customers waiting there, he immediately starts servicing them, but when on his return his queue is empty again, he takes “another vacation” from his main queue to serve somewhere else, and so on. For a survey, see [66].

The interplay of nodes with IVQ and local stability issues has been studied in depth in [172]. Reliability of nodes is not considered there nor is the impact of parameter insecurity addressed.

Parameter insecurity is well-studied in the simulation literature and for details we refer to the excellent overviews [25, 94]. In case the uncertainty in the model parameters is given through deterministic bounds rather than statistical information, computing bounds on the model output due to the perturbations in the input parameters is feasible. Ramesh et al. derive analytic bounds for model output for variations in single parameter value or multiparameter values for Markov models [163]. Parametric sensitivity analysis via derivatives of the output measure with respect input parameters has also been used frequently in this context [164, 36]. Our approach to parameter insecurity is closely related to performability modelling, where performability is the
combined modelling of performance and dependability of fault-tolerant and distributed computer and communications systems, see [181, 86, 87]. Haverkort et. al. [88] compute the mean and variance of the model outputs by the method of parametric sensitivity analysis as well as the quantiles of the distribution of model output (along with mean and variance) by propagating the uncertainty through the model using Monte Carlo sampling. Our approach replaces Monte Carlo sampling by explicit computation of the density of the output. Moreover, we show that the value of risk can be explicitly computed.

3.4 Extended Jackson Networks

In this section, we consider Jackson networks which at some nodes have infinite supply and where some nodes break down randomly and are repaired thereafter. Breakdowns of nodes in standard Jackson networks were investigated in [165] and [155]. It turns out that breakdown of nodes with infinite supply require a more specific regime to control breakdown and repair. The consequences of breakdown of a node are as follows. Whenever a node breaks down,

- service at this node is interrupted, customers (of high as well as of low priority) are frozen there to wait for restart of the service, which is resumed at the point where it was pre-empted,
- no new customers are admitted to enter that node,
- customers who select a broken down node to visit are rerouted according one of the classical rules: stalling, skipping or blocking rs-rd, which will be defined next,
- all these rules, if applicable, are valid for both classes, high and low priority, of customers.

Note, that rerouting does not apply directly to low priority customers, because on departure from a node with infinite supply they are transformed immediately to high priority, and only thereafter necessary rerouting is in force.

3.4.1 The Rerouting Regimes

We follow [165] and distinguish the following three rerouting rules:

**Stalling:** Whenever a node breaks down the service system is frozen, i.e., all arrival processes are interrupted and the service anywhere in the network is stopped. Thus every movement of customers inside the network and arrivals to the network from the outside are stopped until all broken down nodes are repaired again, i.e., if nodes in $I \neq \emptyset$ are broken down then for all $i \in \bar{J}$ the $I$-dependent rates are set to $\lambda_i^I = \mu_i^I = 0$. We assume that the stopped nodes which are in up status are waiting in warm standby, i.e., they can break down although they are stalled. Stalling is applied, for example, in the automotive industry for decreasing variability of the flow of materials. Indeed, stalling prevents that servers continue sending parts to a server that is broken down and thereby prevents piling up inventory.

**Skipping:** Customers are not allowed to enter down nodes and have to skip these nodes. I.e., if the next destination of a customer is a down node, the customer jumps to
this node spending no time there and immediately performs the next jump according to his routing regime until he arrives at a node in up status or leaves the network. Whenever a breakdown of a subset $I \subseteq D$ occurs, customers are rerouted according to the following routing matrix $R^I = (r^I(i, j) : i, j \in \{0\} \cup \bar{J} \setminus I)$:

$$r^I(j, k) = r(j, k) + \sum_{i \in I} r(j, i) r^I(i, k) \text{ for } k, j \in \{0\} \cup \bar{J} \setminus I$$

with

$$r^I(i, k) = r(i, k) + \sum_{l \in I} r(i, l) r^I(l, k) \text{ for } i \in I, k \in \{0\} \cup \bar{J} \setminus I. \quad (3.2)$$

The external arrival rates during a breakdown of $I$ are

$$\lambda^I_j = \lambda_j + \sum_{i \in I} \lambda_i r^I(i, j) \text{ for } j \in \bar{J} \setminus I$$

and $\lambda^I_k = 0$ for $k \in I$. The service intensities are

$$\mu^I_i = \begin{cases} \mu_i, & i \in \bar{J} \setminus I, \\ 0, & \text{otherwise}. \end{cases}$$

**Blocking rs-rd:** Broken down stations are blocked. A customer whose next destination is a down node stays at his present node to obtain another service there. After the repeated service (rs) the customer chooses his next destination anew according to his routing matrix, i.e., a random destination (rd). Whenever a breakdown of a subset $I \subseteq D$ occurs, customers are rerouted according to the following routing matrix $R^I = (r^I(i, j) : i, j \in \{0\} \cup \bar{J} \setminus I)$ with

$$r^I(i, j) = \begin{cases} r(i, j), & i, j \in \{0\} \cup \bar{J} \setminus I, i \neq j, \\ r(i, j) + \sum_{k \in I} r(i, k), & i \in \{0\} \cup \bar{J} \setminus I, \ i = j. \end{cases} \quad (3.3)$$

The external arrival rates during a breakdown of $I$ are $\lambda^I_j = \lambda_j$ for $j \in \bar{J} \setminus I$ and $\lambda^I_k = 0$ otherwise as well as the service intensities are

$$\mu^I_i = \begin{cases} \mu_i, & i \in \bar{J} \setminus I, \\ 0, & \text{otherwise}. \end{cases}$$

With these preparing definitions we summarize our construction. Consider a Jackson network with infinite supply and unreliable nodes and rerouting regime is either stalling, skipping, or blocking rs-rd with the respective rerouting matrices $R^I$. Denote $R^\emptyset := R$ and $N_0 := N \cup \{0\}$. Then the joint availability-queue length process is described by the Markov process $(Y, X) = ((Y(t), X_1(t), ..., X_J(t)) : t \in \mathbb{R}_+)$ on the state space $\mathcal{P}(D) \times N_0^J$ with transition rates matrix $Q = (q(z, z') : z, z' \in \mathcal{P}(D) \times N_0^J)$ defined for all $(n_1, ..., n_J) \in N_0^J, I \subseteq D$ and $i, j \in \bar{J} \setminus I, i \neq j$ as:

$$q(I, n_1, ..., n_i, ..., n_j; I, n_1, ..., n_i + 1, ..., n_j) = \lambda^I_j + \sum_{k \in V \setminus I} \mu^I_k r^I(k, i) 1_{\{n_k = 0\}},$$
where $1_{\{n_k=0\}}$ is an indicator function with 1 if $n_k = 0$ and 0 else,

$$q(I, n_1, ..., n_j; I, n_1, ..., n_i - 1, ..., n_J) = \mu_j^r(r(i, 0))1_{\{n_i>0\}},$$

$$q(I, n_1, ..., n_j; I, n_1, ..., n_i - 1, ..., n_j + 1, ..., n_J) = \mu_j^r(r(j, j))1_{\{n_i>0\}},$$

$$q(I, n_1, ..., n_j; H, n_1, ..., n_J) = \alpha(I, H) = \frac{A(H)}{A(I)}, \quad I \subset H \subset D,$n

$$q(I, n_1, ..., n_J; K, n_1, ..., n_J) = \beta(I, K) = \frac{B(I)}{B(K)}, \quad K \subset I \subset D,$n

$$q(I, n_1, ..., n_J; I, n_1, ..., n_J) = - \sum_{i \in J \setminus I} \lambda_i^I - \sum_{i \in J \setminus I} \sum_{k \in V \setminus I} \mu_k^r(r(k, i))1_{\{n_k=0\}} - \sum_{I \subset H \subset D} \alpha(I, H) - \sum_{K \subset I \subset D} \beta(I, K),$$

and $q(z; z') = 0$ otherwise for $z \neq z'$.

The following theorem yields a characterization of the departure streams from nodes.

**Theorem 3.1.** With the above definitions:

(i) If at time $t$ all nodes are up, i.e., $Y(t) = \emptyset$, then the departure streams from node $j \in V$ is a Poisson process with rate $\mu_j$. Thus the departure stream from $j \in V$ to $i \in J$ is Poisson with rate $\mu_j r(j, i)$.

(ii) Whenever nodes in $I \neq \emptyset$ are broken down and either skipping or blocking rs-rd is in force, the departure stream of node $j \in V \setminus I$ with infinite supply in up-status is Poisson with rate $\mu_j$. The departure stream from $j \in V \setminus I$ to $i \in J \setminus I$ is Poisson with rate $\mu_j r(j, i)$, where $r(j, i)$ is determined by the rerouting regime in force.

In the case of stalling, all Poisson arrival streams stop whenever a breakdown occurs ($I \neq \emptyset$), and the Poisson streams are reactivated when all nodes recur to the up-status.

**Proof.** Consider the network in its different availability states of nodes:

**Case I:** All nodes are in up-status ($I = \emptyset$), all departure times from node $j \in V$

- are independent of the state of the node due to the infinite supply,
- and therefore have independent and identically exponentially distributed interdeparture times with rate $\mu_j$, the service rate of the node.

Thus, the departure stream of node $j \in V$ is a Poisson process with rate $\mu_j$. Hence for node $i \in J$, the arrival stream from node $j \in V$ is a Poisson process with rate $\mu_j r(j, i)$, because a portion of $r(j, i)$ of the departure stream is directed to node $i \in J$.

**Case II:** Nodes in $I \neq \emptyset$ are broken down and rerouting is according to blocking rs-rd or skipping, all departure times from node $j \in V \setminus I$

- are independent of the state of the node due to the infinite supply,
- and therefore have independent and identically exponentially distributed interdeparture times with rate $\mu_j$, the service rate of the node.
Therefore the departure stream of node $j \in V$ is a Poisson process with rate $\mu_j$. Thus, for node $i \in \tilde{J} \setminus I$, the arrival stream from node $j \in V \setminus I$ is a Poisson process with rate $\mu_j r^t(j, i)$, because a portion of $r^t(j, i)$ of the departure stream is directed to node $i \in \tilde{J} \setminus I$. Whenever a node with infinite supply breaks down, its Poisson departure stream is interrupted until the node is repaired.

**Case III:** Nodes in $I \neq \emptyset$ are broken down, under stalling all network processes - except breakdown and repair processes - are frozen, i.e., no service is provided in the network and there is no arrival stream. Thus, as long as not all nodes are in up status, there are no customer flows in the network.

### 3.4.2 Extended Traffic Equations

Different traffic equations that are required for analysis of the long-term behaviour are provided in the subsequent definitions. We start with providing the traffic equations for networks without breakdowns and repairs and furthermore formalize the notion of stable nodes.

**Definition 3.2** (Definition 5 from [172]). The general traffic equations for Jackson networks with infinite supply but no breakdowns and repairs are

$$
\eta_i = \lambda_i + \sum_{j \in W} \min(\eta_j, \mu_j) r(j, i) + \sum_{j \in V} \mu_j r(j, i), \quad i \in \tilde{J}.
$$

(3.4)

A node $i$ is stable if $\eta_i$ determined by (3.4) is strictly less than $\mu_i$, otherwise the node is unstable.

The above traffic equations are motivated by the following considerations as given in [172]:

1. For node $j \in V$ with infinite supply the output rate of high priority jobs is $\mu_j$, which usually is not the overall arrival rate $\eta_j$ at node $j$.

2. For a stable node $j \in W$ (without infinite supply) the overall arrival rate $\eta_j$ is the maximal departure rate as well, which can be met by the node because $\mu_j > \eta_j$.

3. For an unstable node $j \in W$ the overall arrival rate $\eta_j$ in general cannot be met by the node’s capacity, because it can maximally process at rate $\mu_j$.

4. Arguments (ii) and (iii) lead to the departure rates $\min(\eta_j, \mu_j)$ from nodes $j \in W$.

Furthermore, in Lemma 6 of [172] it is shown that the traffic equations from Definition 3.2 have a unique solution (because of the irreducibility assumption of $R$).

It is worth noting that for networks where all nodes in $W$ are stable the traffic equations in Definition 3.2 reduce to the (standard) traffic equations of a Jackson network with infinite supply [188]

$$
\eta_i = \lambda_i + \sum_{j \in W} \eta_j r(j, i) + \sum_{j \in V} \mu_j r(j, i), \quad i \in \tilde{J}.
$$

(3.5)

In the following we present the versions of the traffic equations which are appropriate for the specific breakdown regimes in case all nodes without infinite supply are stable, i.e., $W \cap U = \emptyset$.
Definition 3.3. The (standard) traffic equations for unreliable Jackson networks with infinite supply and $W \cap U = \emptyset$ are as follows:

(i) In case of stalling,

\[ \eta_i = \lambda_i + \sum_{j \in W} \eta_j r(j, i) + \sum_{j \in V} \mu_j r(j, i), \quad i \in \tilde{J}, \]

as long as all nodes are in up status ($I = \emptyset$). Otherwise $\eta_i^I = 0$ for all $i \in \tilde{J}$.

(ii) In case of rerouting according to blocking rs-rd or skipping,

\[ \eta_i^I = \lambda_i^I + \sum_{j \in W \setminus I} \eta_j^I r(j, i) + \sum_{j \in V \setminus I} \mu_j^I r(j, i), \quad i \in \tilde{J} \setminus I, \]  

(3.6)

for all $I \subseteq D$ and $\eta_j^I = 0$ for all $j \in I$.

By noting that for all $I \subseteq D$ the routing matrix $R^I$ (for blocking rs-rd and skipping) remains irreducible for nodes from $\{0\} \cup \tilde{J} \setminus I$, it follows from the proof of Lemma 4 in [172] (by disregarding nodes from $I$) that the traffic equations (3.6) have a unique solution $\eta^I = (\eta_i^I : i \in \tilde{J} \setminus I), \forall I \subseteq D$ (for blocking rs-rd and skipping).

The traffic equations from Definition 3.3 for some $I \subseteq D$ remain valid only as long as the availability status is unchanged. Whenever the availability status of the system changes, the traffic equations are adapted according to the new set of broken down nodes. Thus each traffic equation (3.6) may have different solutions for different $I$. The following two lemmas show under which constraints the solution of the traffic equation (3.6) remains the same on $\tilde{J} \setminus I$ for all $I \subseteq D$ in case of blocking rs-rd and skipping regimes, respectively.

Lemma 3.1. Consider a Jackson network where nodes in $D \subseteq \tilde{J}$ are unreliable and nodes in $V \subseteq \tilde{J}$ have an infinite supply of work. $W := \tilde{J} \setminus V$ is the set of nodes without infinite supply. Let for all nodes $i \in W$ without infinite supply hold $\eta_i < \mu_i$, where $(\eta_i : i \in \tilde{J})$ is the unique solution of (3.5). In case of breakdowns of nodes we assume that customers are rerouted according to the blocking rs-rd regime.

(i) If the following reversibility constraints hold:

\[ \eta_r(i, j) = \eta_j r(j, i) \quad \forall i, j \in W, \]  

(3.7)

\[ \eta_r(i, j) = \mu_j r(j, i) \quad \forall i \in W, j \in V, \]  

(3.8)

Then for all nodes $i \in W \setminus I$ the solution $\eta_i^I$ of the traffic equation (3.6) equals $\eta_i$ for all $I \subseteq D$.

(ii) Let (3.7) and (3.8) hold. If we additionally require the reversibility constraint

\[ \mu_r(i, j) = \mu_j r(j, i) \quad \forall i, j \in V, \]  

(3.9)

Then for the solution of (3.6) holds $\eta_i^I = \eta_i$ for all $i \in V \setminus I$ and all $I \subseteq D$, as well.
Proof. (i): We will show that the unique solution (η_i : i ∈ J) of (3.5) is also a solution for (3.6). In particular, for any \( I \subseteq D \) and \( \forall i \in W \setminus I \)

\[
\lambda_i^I + \sum_{j \in W \setminus I} \eta_j r^I(j, i) + \sum_{j \in V \setminus I} \mu_j r^I(j, i)
= \lambda_i + \eta_i \left( r(i, i) + \sum_{k \in I} r(i, k) \right) + \sum_{j \in W \setminus I, j \neq i} \eta_j r(j, i) + \sum_{j \in V \setminus I} \mu_j r(j, i)
= \lambda_i + \sum_{k \in I} \eta_k r(k, i) + \sum_{k \in I} \eta_k r(i, k) + \sum_{j \in W \setminus I} \eta_j r(j, i) + \sum_{j \in V \setminus I} \mu_j r(j, i)
= \lambda_i + \sum_{j \in W} \eta_j r(j, i) + \sum_{j \in V} \mu_j r(j, i)
\]

Because of the uniqueness of solution (η_i^I : i ∈ J \setminus I) in (3.6), \( \forall I \subseteq D \) (see also the comment after Definition 3.3), this proves (i).

(ii): For any \( I \subseteq D \) and \( \forall i \in V \setminus I \):

\[
\eta_i^I \overset{(3.6)}{=} \lambda_i^I + \sum_{j \in W \setminus I} \eta_j^I r^I(j, i) + \sum_{j \in V \setminus I} \mu_j r^I(j, i)
\overset{(3.3)}{=} \lambda_i + \sum_{j \in W \setminus I} \eta_j r(j, i) + \mu_i \left( r(i, i) + \sum_{k \in I} r(i, k) \right) + \sum_{j \in V \setminus I, j \neq i} \mu_j r(j, i)
\overset{(3.8)}{=} \lambda_i + \sum_{k \in I} \eta_k r(k, i) + \sum_{k \in I} \mu_k r(k, i) + \sum_{j \in V} \mu_j r(j, i)
\overset{(3.9)}{=} \lambda_i + \sum_{j \in W} \eta_j r(j, i) + \sum_{j \in V} \mu_j r(j, i)
\]

Therefore, η_i^I = η_i.

\[\square\]

Remark 3.1. Due to the infinite supply, the reversibility constraints (3.7), (3.8) and (3.9) are different from the classical reversibility constraints which are the local balance equations of the routing process. But the interpretation of (3.7), (3.8) and (3.9) is the same: the departure rate from i in the direction j equals the departure rate from j in the direction i.

For rerouting in order to skip broken down nodes, i.e., the skipping regime applies, we assume that the unreliable nodes in V are rate stable according to [126] on p. 76, i.e., these nodes have equal input and output rates. Because such a node \( i \in V \cap D \) has precisely an output rate (of priority jobs) \( \mu_i \) due to the infinite supply, it is on average, fully loaded already by customers of pre-emptive priority. So, intuitively, there are only a few low priority customers (from the infinite supply) served.

Lemma 3.2. Let for all nodes \( i \in W \) without infinite supply hold \( \eta_i < \mu_i \) where (\( \eta_i : i \in J \)) is the unique solution of (3.5). In case of breakdowns of nodes we assume that customers are rerouted according to the skipping regime. Let the following constraint (rate stability) hold:

\[
\eta_i = \mu_i, \quad \forall i \in V \cap D.
\]

\[\text{(3.10)}\]
Then for all nodes \( i \in \bar{J} \setminus I \) the solution \( \eta^I_i \) of the traffic equation (3.6) for all \( I \subseteq D \) equals \( \eta_i \).

**Proof.** We will show that the unique solution \((\eta_i : i \in \bar{J})\) of (3.5) is also a solution for (3.6). In particular, for any \( I \subseteq D \) and \( \forall i \in W \setminus I \)

\[
\lambda^I_i + \sum_{j \in W \setminus I} \eta_j r^I(j, i) + \sum_{j \in V \setminus I} \mu_j r^I(j, i)
= \lambda_i + \sum_{k \in I} \lambda_k r^I(k, i) + \sum_{j \in W \setminus I} \eta_j \left( r(j, i) + \sum_{k \in I} r(j, k) r^I(k, i) \right)
+ \sum_{j \in V \setminus I} \mu_j \left( r(j, i) + \sum_{k \in I} r(j, k) r^I(k, i) \right)
= \lambda_i + \sum_{j \in W \setminus I} \eta_j r^I(j, i) + \sum_{j \in W \setminus I} \sum_{k \in I} \mu_j r^I(j, k) + \lambda_i \sum_{j \in W \setminus I} \eta_j r^I(j, k) + \sum_{j \in W \setminus I} \eta_j r^I(j, k)
= \eta_i - \sum_{j \in I \setminus W} \eta_j r^I(j, i) - \sum_{j \in I \setminus V} \mu_j r^I(j, i) + \sum_{j \in I \setminus W} \eta_j r^I(j, k) - \sum_{k \in I} \sum_{j \in I \setminus W} \eta_j r^I(j, k)
- \sum_{k \in I} r^I(k, i) \sum_{j \in I \setminus W} \mu_j r^I(j, k)
= \eta_i - \sum_{j \in I \setminus W} \eta_j \left( r(j, i) + \sum_{k \in I} r(j, k) r^I(k, i) \right) + \sum_{k \in I} \eta_k r^I(k, i)
- \sum_{j \in I \setminus W} \mu_j \left( r(j, i) + \sum_{k \in I} r(j, k) r^I(k, i) \right)
= \eta_i + \sum_{k \in I \setminus W} \eta_k r^I(k, i) - \sum_{j \in I \setminus W} \eta_j r^I(j, i) + \sum_{k \in I \setminus W} \eta_k r^I(k, i) - \sum_{j \in I \setminus W} \mu_j r^I(j, i)
= \eta_i + \sum_{k \in I \setminus W} \left( \eta_k - \mu_k \right) r^I(k, i) = \eta_i.
\]

Because of the uniqueness of solution \((\eta^I_i : i \in \bar{J} \setminus I)\) in (3.6), \( \forall I \subseteq D \) (see also the comment after Definition 3.3), this proves that \( \eta^I_i = \eta_i \) holds for all \( j \in W \setminus I \) and all \( I \subseteq D \). Using this result it follows for all \( i \in V \setminus I \) and \( I \subseteq D \):

\[
\eta^I_i = \lambda^I_i + \sum_{j \in W \setminus I} \eta^I_j r^I(j, i) + \sum_{j \in V \setminus I} \mu_j r^I(j, i) = \lambda^I_i + \sum_{j \in W \setminus I} \eta^I_j r^I(j, i) + \sum_{j \in V \setminus I} \mu_j r^I(j, i)
\]

which is the very left side of (3.11) with \( i \in V \setminus I \). The above computations in (3.11) are valid for all \( i \in \bar{J} \setminus I \), hence it follows \( \eta^I_i = \eta_i \) for all \( i \in V \setminus I \) and \( I \subseteq D \), too.

We illustrate the adjusted traffic equations with the example presented in Section 3.2, where nodes in \( D = \{3, 5\} \) are unreliable. Then the condition (3.10) requires \( \eta_5 = \mu_5 \) because node 5 is not stable and (3.4) is the relevant traffic equation in case of no breakdowns. Due to the feed forward structure of the network we can evaluate the traffic rates directly.
3.4. Extended Jackson Networks

Example 3.1. For \( I = \emptyset \), \( \eta \) is just the solution of the standard traffic equation (3.4) given by:

\[
\begin{align*}
\eta_1 &= \lambda_1, \quad \eta_2 = \lambda_2, \quad \eta_3 = \mu_1, \quad \eta_4 = r(3, 4)\mu_1 + r(2, 4)\lambda_2, \\
\eta_5 &= r(3, 5)\mu_1, \quad \eta_6 = r(3, 4)\mu_1 + r(2, 4)\lambda_2 + r(5, 6)\mu_5, \quad \eta_7 = r(2, 7)\lambda_2.
\end{align*}
\]

Now suppose nodes in \( D = \{3, 5\} \) are down and consider stalling, then \( \eta_i = 0 \) for all \( i \). In case of skipping, we obtain \( \eta_i^f = \eta_i \), for \( i \in \{1, 2, 4, 6, 7\} \), and \( \eta_j^f = 0 \), for \( j = 3, 5 \). Eventually, blocking rs-rd can not be implemented in this network if \( \eta_i^f = \eta_i \), for \( i \in \{1, 2, 4, 6, 7\} \), is required because the routing chain is not reversible.

Blocking rs-rd does not apply to the system put forward in Example 3.1. In the following we will present two examples of networks that do meet the requirements. One will be a network having a linear topology and the other one having a star-shaped topology.

Example 3.2. (Two-way tandem network) Consider a network with \( J = \{1, 2, 3\} \), \( V = \{2\} \), \( W = \{1, 3\} \), and \( D = V \), i.e., the infinite supply node is prone to failure. Routing is given by

\[
r(1, 2) = a, \quad r(1, 0) = 1 - a, \quad r(2, 3) = b, \quad r(2, 1) = 1 - b, \quad r(3, 0) = c, \quad r(3, 2) = 1 - c,
\]

and

\[
r(0, 1) = \frac{\lambda_1}{\lambda_1 + \lambda_3}, \quad r(0, 3) = \frac{\lambda_3}{\lambda_1 + \lambda_3},
\]

for \( 0 < a, b < 1 \) and \( \lambda_i > 0 \), \( i = 1, 2 \). The network can be visualized as a two-way tandem of three nodes, see Figure 3.2. The infinite supply is depicted by a dashed arrow pointing to server 2, and a bold circle is shown to indicate that node 2 is prone to failure. Note that by incorporating node 0 into the network, the linear topology is transformed into a ring.

For ease of analysis we parameterize the model in the following way. We let \( \lambda_1 = (1 - a)t \), for \( t > 0 \), and \( \lambda_3 = at \), and \( b = 1 - c \). The service rates are set to

\[
\mu_1 > t, \quad \mu_2 = \frac{a}{c} \quad \text{and} \quad \mu_3 > \frac{a}{c}.
\]

The standard traffic equations (3.4) then have the solution

\[
\eta_1 = t, \quad \eta_2 = \eta_3 = \frac{a}{c}.
\]
Note that this implies that \( U = \{2\} \), i.e., the set of unstable nodes consists only of node 2.

With this choice of parameters it can be seen after some tedious algebra that the network satisfies the reversibility conditions put forward in Lemma 3.1 and rate stability in Lemma 3.2 thus all three blocking disciplines apply. Recall that \( D = \{2\} \). Hence, \( \emptyset, \{2\} \) are the only two possible breakdown scenarios and with the breakdown and repair rates from the motivating example in Section 3.2 we have

\[
\pi(\emptyset) = \frac{\rho_2}{\rho_2 + \tau_2} \quad \text{and} \quad \pi(\{2\}) = \frac{\tau_2}{\rho_2 + \tau_2}.
\]

\( \square \)

**Example 3.3. (Star-shaped network)** Consider a network with \( J = \{1, 2, \ldots, 6\} \), \( V = \{2, 3, 4\} \), \( W = \{1, 5, 6\} \), and \( D = V \), i.e., all infinite supply nodes are prone to failure. Jobs arrive from the outside with rate \( \lambda \) to node 1. From node 1 they go with probability \( r/5 \) to any of the nodes 2 to 6, for \( r \in (0, 1) \). After finishing service at node \( i = 2, \ldots, 6 \), jobs are sent back to the central node 1. From there they leave the system with probability \( 1 - r \), or are sent back to servers in the set \( \{2, \ldots, 6\} \) according to the routing scheme described above. The network can be visualized as a star-shaped network with one central node and five nodes in the periphery, see Figure 3.3. The infinite supply is depicted by a dashed arrow pointing to servers 2, 3, 4, and the nodes that are prone to failure are depicted by bold circles. The traffic equations are given by

\[
\eta_1 = \lambda + \mu_2 + \mu_3 + \mu_4 + \eta_5 + \eta_6,
\]

where \( \eta_5 = \eta_6 = r\eta_1/5 \), provided that nodes 5, 6 are stable. For **blocking rs-rd** and **skipping** to be applicable, we let

\[
\eta_i = \frac{r}{5}\eta_1 = \mu_i, \quad 2 \leq i \leq 4,
\]

\( \square \)
which implies \( \eta_i = \lambda/(1-r) \), and thus
\[
\frac{r \lambda}{5(1-r)} = \mu_i = \eta_i, \quad 2 \leq i \leq 4.
\]

Eventually, we let
\[
\frac{r \lambda}{5(1-r)} < \mu_5, \mu_6,
\]
in order to let 5, 6 be stable nodes. Indeed, this choice implies \( \eta_i < \mu_i, \quad i = 5, 6 \). The above conditions imply that the reversibility conditions put forward in Lemma 3.1 and the rate stability condition put forward in Lemma 3.2 are satisfied for the star-shaped network.

Recall that \( D = \{2, 3, 4\} \). Hence, \( \emptyset, \{2\}, \{3\}, \{4\}, \{2, 3\}, \{3, 4\}, \{2, 4\} \) and \( D \) are the possible breakdown scenarios and with the breakdown and repair rates from the motivating example in Section 3.2 we have
\[
\pi(\emptyset) = \left(1 + \sum_{i=2}^{4} \frac{\tau_i}{\rho_i} \right) \left( \frac{\tau_2 + \tau_3}{2 \min(\rho_2, \rho_3)} + \frac{\tau_2 + \tau_4}{2 \min(\rho_2, \rho_4)} + \frac{\tau_3 + \tau_4}{2 \min(\rho_3, \rho_4)} + \frac{\tau_2 + \tau_3 + \tau_4}{3 \min(\rho_2, \rho_3, \rho_4)} \right)^{-1}
\]
and
\[
\pi\{i\} = \pi(\emptyset) \frac{\tau_i}{\rho_i},
\]
for \( i = 2, 3, 4 \),
\[
\pi\{i,j\} = \pi(\emptyset) \frac{\tau_i + \tau_j}{2 \min(\rho_i, \rho_j)},
\]
for \( 2 \leq i,j \leq 4 \), with \( i \neq j \), and
\[
\pi(D) = \pi(\emptyset) \frac{\tau_2 + \tau_3 + \tau_4}{3 \min(\rho_2, \rho_3, \rho_4)}.
\]

\subsection{3.4.3 Long-Term Behaviour}

In this subsection we will study the long-term behaviour of extended Jackson networks. In this context, two terms to be considered are the limiting distributions and stationary distributions, respectively.

**Definition 3.4** (Based on Definition 1.3 from [172]). Let \( (Y, X) = (Y(t), X_1(t), \ldots, X_J(t) : t \in \mathbb{R}_+) \) be a homogeneous Markov process with discrete state space \( \mathcal{P}(D) \times \mathbb{N}_0^J \). \((Y, X)\) has a limiting distribution if
\[
\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i : i \in \bar{J}) \text{ exists for all } (I, n_i : i \in \bar{J}) \in \mathcal{P}(D) \times \mathbb{N}_0^J
\]
and
\[
\sum_{(I,n_i;i\in\bar{J})\in\mathcal{P}(D)\times\mathbb{N}_0^J} \lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i : i \in \bar{J}) = 1.
\]

**Definition 3.5** (Based on Lemma 1.5 from [172]). Let \( (Y, X) = (Y(t), X_1(t), \ldots, X_J(t) : t \in \mathbb{R}_+) \) be a homogeneous Markov process with discrete state space \( \mathcal{P}(D) \times \mathbb{N}_0^J \) with transition rates matrix \( Q = (q(z, z') : z, z' \in \mathcal{P}(D) \times \mathbb{N}_0^J) \) which is irreducible and positive recurrent, i.e., ergodic. A probability measure \( \pi \) is a stationary distribution for \((X, Y)\) if and only if it solves the global balance equation
\[
\pi Q = 0.
\]
A limiting distribution for a Markov chain is always a stationary distribution. If a Markov process is ergodic then its stationary distribution is the limiting distribution, see, e.g., p. 40 from [168].

**Theorem 3.2.** Let \( W \cap U = \emptyset \), so all nodes without infinite supply are stable. Denote by \( \eta = (\eta_1, ..., \eta_J) \) the unique solution of the traffic equations (3.5).

Under the **stalling** regime, it holds that:

(i) For nodes without infinite supply, the joint marginal limiting distribution is:

\[
\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i : i \in W) = \left( \sum_{K \subseteq D} \frac{A(K)}{B(K)} \right)^{-1} \frac{A(I)}{B(I)} \prod_{i \in W} \left( 1 - \frac{\eta_i}{\mu_i} \right) \left( \frac{\eta_i}{\mu_i} \right)^{n_i},
\]

(3.13)

for all \( I \subseteq D \) and all \( (n_i : i \in W) \in \mathbb{N}_0^W \), and this is a stationary distribution on \( W \) as well.

(ii) If the global network process is started with an initial distribution which has the marginal (3.13) on \( W \), the arrival stream from \( i \in W \) to \( j \in V \) is Poisson with rate \( \eta_i r(i, j) \) whenever all nodes are in up status. And these streams are independent given that the nodes are up.

(iii) If the global network process is started with an initial distribution which has the marginal (3.13) on \( W \), then the marginal limiting distribution for a stable node \( i \in V \) with \( r(i, i) = 0 \) is:

\[
\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i) = \left( \sum_{K \subseteq D} \frac{A(K)}{B(K)} \right)^{-1} \frac{A(I)}{B(I)} \left( 1 - \frac{\eta_i}{\mu_i} \right) \left( \frac{\eta_i}{\mu_i} \right)^{n_i},
\]

(3.14)

for all \( I \subseteq D \) and all \( n_i \in \mathbb{N}_0 \), if and only if \( \eta_i < \mu_i \), and this is a one-dimensional stationary distribution as well.

If \( \eta_i \geq \mu_i \) for node \( i \in V \), then for its limiting probability holds for all \( I \subseteq D \) and all \( n_i \in \mathbb{N}_0 \):

\[
\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i) = 0,
\]

(3.15)

if the global network process is started with initial distribution which has marginal (3.13) on \( W \).

The proof of Theorem 3.2 can be found in Appendix B and also the proof of Theorem 3.3 below can be found in that appendix.

**Theorem 3.3.** Assume that \( W \cap U = \emptyset \), so all nodes without infinite supply are stable. Denote by \( \eta = (\eta_1, ..., \eta_J) \) the unique solution of the traffic equations (3.5). In case of breakdowns customers are rerouted according to the blocking rs-rd regime or the skipping regime. If blocking rs-rd is in force we require the reversibility-constraints (3.7) and (3.8). If skipping is in force, let (3.10) hold. Then for nodes without infinite supply the joint marginal limiting distribution is given by:

\[
\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i : i \in W) = \left( \sum_{K \subseteq D} \frac{A(K)}{B(K)} \right)^{-1} \frac{A(I)}{B(I)} \prod_{i \in W} \left( 1 - \frac{\eta_i}{\mu_i} \right) \left( \frac{\eta_i}{\mu_i} \right)^{n_i},
\]

(3.16)
for all $I \subseteq D$ and all $(n_i : i \in W) \in \mathbb{N}_0^{\lvert W \rvert}$, and this is a stationary distribution on $W$ as well.

The results of Goodmann and Massey [79] are on classical Jackson networks where some nodes are not stable. They prove a product-form limiting distribution for the stable subnetwork, but there is no such result for the stationary distribution, i.e. for the finite time horizon. The reason is that the exploding unstable nodes will influence the stable part of the network and over any finite (transient) time horizon $[0,t]$ the departure streams from the unstable nodes fail to be of Poisson type. Put differently, only the limiting distribution is known. Fortunately, the proofs of Theorem 3.2(i) and (iii) and of Theorem 3.3 allow for establishing the following result on the transient phase of the network process.

**Corollary 3.1.** Under the conditions put forward in Theorem 3.2 and Theorem 3.3, the process $(Y, X_W) := (Y, X_i : i \in W)$ is an ergodic homogeneous Markov process of its own. If, in addition, for $i \in V$ it holds that $\eta_i < \mu_i$, then the process $(Y, X_i)$ is an ergodic homogeneous Markov process of its own for $i \in V$.

**Remark 3.2.** In the setting of Theorem 3.3 a statement as in Theorem 3.2(iii) cannot be proved with the methods used here. This is due to the properties of the rerouting regimes skipping and blocking rs-rd. Whenever nodes in $I \neq \emptyset$ are down, immediate feedback may emerge even at nodes in $V$ with $r(i,i) = 0$. If $i \in V \setminus I$ and $r(i,j) > 0$ for at least one $j \in I$ holds, then $r^I(i,i) > 0$ may occur.

On the other hand, if the network’s topology prevents occurrence of feedback by skipping or rs-rd regime in case of breakdown it is possible to prove a counterpart to Theorem 3.2(iii) in the setting of Theorem 3.3.

Our motivating example from Section 3.2 in Figure 3.1 is a feed forward network according to the following definition. Feed forward networks are an important subclass of Jackson networks.

**Definition 3.6.** A network with node set $\tilde{J}$ with $\lvert \tilde{J} \rvert = J$ is a feed forward network if there exists an enumeration $\tilde{J} = \{1, 2, \ldots, J\}$ of the nodes such that

$$r(i,j) > 0 \implies i < j$$

holds.

Note that a feed forward network can not be reversible and therefore in case of breakdowns we must recur to stalling or skipping as rerouting scheme, i.e., the reversibility constraints of Lemma 3.1 do not hold, so that the results for breaking rs-rd cannot be applied. The following property of feed forward networks is intuitive.
Lemma 3.3. If in a feed forward network with node set $\tilde{J} = \{1, 2, \ldots, J\}$ a subset $\emptyset \subseteq I \subseteq \tilde{J}$ of nodes is down and either skipping or stalling is applied as rerouting scheme, then it holds that
$$r^I(i, j) > 0 \implies i < j,$$
and therefore there is no immediate feedback at all nodes.

Theorem 3.4. Consider a feed forward network with node set $\tilde{J} = \{1, 2, \ldots, J\}$. Assume that $W \cap U = \emptyset$, so all nodes without infinite supply are stable. Denote by $\eta = (\eta_1, \ldots, \eta_J)$ the unique solution of the traffic equations (3.5). In case of breakdowns customers are rerouted according to the skipping regime and assume that (3.10) holds.

If the global network process is started with an initial distribution which has the marginal (3.13) on $W$, then the marginal limiting distribution for a stable node $i \in V$ is:
$$\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i) = \left( \sum_{K \subseteq D} A(K)B(K) \right)^{-1} \frac{A(I)B(I)}{B(I)} \left(1 - \frac{\eta_i}{\mu_i}\right)^{n_i},$$
for all $I \subseteq D$ and all $n_i \in \mathbb{N}_0$, if and only if $\eta_i < \mu_i$, and this is a one-dimensional stationary distribution as well.
If $\eta_i \geq \mu_i$ for node $i \in V$, then for its limiting probability holds for all $I \subseteq D$ and all $n_i \in \mathbb{N}_0$:
$$\lim_{t \to \infty} P(Y(t) = I, X_i(t) = n_i) = 0,$$
if again the global network process is started with initial distribution which has marginal (3.13) on $W$.

The proof of this theorem is analogous to that of Theorem 3.2, part (iii) with the use of Lemma 3.3.

3.4.4 Availability and Performance Measures

Evaluating performance metrics is possible due to explicit access to the joint distribution of availability and (some of the) queue lengths. Standard performance evaluation requires ergodicity of the underlying Markov processes which allows to approximate long-term average cost functions by integrals of the cost function under the stationary distribution. Unfortunately, for the networks with unstable nodes that are non-ergodic, the stationary distribution fails to exist. Our framework overcomes this restriction and allows us to investigate even non-ergodic networks across subnetworks where stabilization in the long-run occurs. As stated in Corollary 3.1 some important subnetworks of stable nodes can be considered as networks of their own. Therefore for these parts we can extend the traditional analysis directly, as will be seen below. But we emphasize that even if there exists no equilibrium on the stable subnetworks, performance analysis for long-term averages of cost functions is possible via integrals of the cost function under the limiting distribution on stable nodes, respective subnetworks. For details we refer to Section 4.2 in [155] and Section 4.6.4 in [154]. In the following, we will state our results for the setting put forward in Corollary 3.1.
Due to the load-independent breakdown and repair rates, the availability process $Y$ in Theorem 3.2 and Theorem 3.3 is an ergodic Markov process of its own with unique limiting and stationary distribution

$$\pi(I) = \left( \sum_{K \subseteq D} \frac{A(K)}{B(K)} \right)^{-1} \cdot \frac{A(I)}{B(I)}, \quad \forall I \subseteq D. \quad (3.17)$$

From this the stationary (time) point availability (PA) of a Jackson network with infinite supply and unreliable nodes (or subnetworks thereof) may be computed similar to [165, p.185] as

$$PA(H)(t) := \sum_{K \subseteq D \setminus H} \pi(K), \quad \text{for} \ H \subseteq D, t \geq 0$$

where $\pi(I)$ is the probability that exactly the nodes in $I \subseteq D$ are under repair, given by (3.17).

In light of Corollary 3.1, let the unique limiting and stationary distribution of $(Y, X_W)$ be denoted by

$$\pi(I, n_i : i \in W), \quad \text{for} \ I \subseteq D \text{ and } (n_i : i \in W) \in \mathbb{N}_0^{\lvert W \rvert}.$$  

In the following we provide an overview of the main performance characteristics for nodes from the stable subnetwork.

(i) Under **stalling**, the stationary throughput at node $i \in W$ ($\text{TH}_i$) is

$$\text{TH}_i := \sum_{(I,n_j : j \in W) \in \mathcal{P}(D) \times \mathbb{N}_0^{\lvert W \rvert}} \pi(I, n_j : j \in W) \cdot \mu_i \cdot \mathbb{1}_{\{I=\emptyset\}} \cdot \mathbb{1}_{\{n_i>0\}} = \eta_i \cdot \pi(\emptyset),$$

the mean asymptotic queue length at node $i$ ($\text{MAQL}_i$) is

$$\text{MAQL}_i := \sum_{(I,n_j : j \in W) \in \mathcal{P}(D) \times \mathbb{N}_0^{\lvert W \rvert}} n_i \cdot \pi(I, n_j : j \in W) = \frac{\eta_i}{\mu_i} \left( 1 - \frac{\eta_i}{\mu_i} \right)^{-1}.$$  

Evoking Little’s law, see [174, 175], the mean waiting time at node $i$ ($\text{MWT}_i$) follows easily. In particular, Little’s law in our context states that

$$\text{MAQL}_i = \eta_i \cdot \pi(\emptyset) \cdot \text{MWT}_i$$

which gives

$$\text{MWT}_i = \frac{1}{(\mu_i - \eta_i) \pi(\emptyset)}.$$  

(ii) Under **blocking rs-rd** and **skipping**, the stationary throughput at a node $i \in W$ is

$$\text{TH}_i = \eta_i \cdot \sum_{I \subseteq D, i \notin I} \pi(I),$$
the mean queue length is as under (i). Little’s law under blocking s-rd and skipping states that

\[ \text{MAQL}_i = \eta_i \cdot \sum_{I \subseteq D, i \not\in I} \pi(I) \cdot \text{MWT}_i \]

so that the mean waiting time

\[ \text{MWT}_i = \frac{1}{(\mu_i - \eta_i) \sum_{I \subseteq D, i \not\in I} \pi(I)}. \]

Indeed, following Theorem 3.2, Theorem 3.3 and Theorem 3.4, the asymptotic mean queue length at a stable node can be computed like in the standard Jackson network case.

### 3.5 Robustness Analysis

In modelling parameter insecurity the choice of the distribution is of importance and one typically chooses a particular distribution based on, possible incomplete, knowledge that is available. For example, if the mean and the variance are known, and if, in addition, we know that the parameter may take values in \( \mathbb{R} \), the most general distribution is the normal distribution, where “most general” refers to the fact that this distribution maximizes the entropy. On the other hand, when, due to expert knowledge, it is known that the parameter falls into an interval, say, \([a, b]\), then the uniform distribution on \([a, b]\) is the entropy maximizing distribution; see, for example, [128]. Alternatively, there may be statistical knowledge available on \( \theta \) based on measurements. Then, the distribution of the statistic used for estimating for \( \theta \) is a natural candidate for the distribution of \( \theta \).

Let \( f_\tau \) denote the density of the breakdown rate \( \tau \) and let \( h(\tau) \) denote some reward function. Think, for example, of \( h \) as the stationary throughput at node \( i \). Provided \( h \) is invertible and the inverse is differentiable with respect to the throughput

\[ g(y) = f_\tau(h^{-1}(y)) \left| \frac{d}{dy} \left( h^{-1}(y) \right) \right| \]

yields the density of the stationary throughput. Based on the distributional assumptions or statistical information comprised in \( f_\tau \), one may, as is common practice in applied probability, take the expected value of \( \tau \), denoted by \( \mu_\tau = \int y f_\tau(y) dy \), as a noise-free approximation of \( \tau \) and subsequently \( h(\mu_\tau) \) as output for the throughput. Since \( h(\mu_\tau) \) is typically not close to \( \mathbb{E}[h(\tau)] \), simply using \( \mu_\tau \) instead of \( \tau \) falls short of bringing the risk incurred by the insecurity on \( \tau \) to light. We argue that the model risk incurred by \( \tau \) is best described by means of the value at risk of \( h(\tau) \), denoted in short by \( \text{VaR}(\alpha) \), where

\[ \text{VaR}(\alpha) = q \text{ if and only if } G^{-1}(\alpha) = q, \]

where \( G(\cdot) \) denotes the cumulative distribution function of \( h(\tau) \), which is, for ease of presentation, assumed to be continuous and invertible. The potential misspecification at an \( \alpha \) probability level is thus \( h(\mu_\tau) - \text{VaR}(\alpha) \). Note that for the throughput we want to hedge against the risk of low values, so we use the \( \alpha \)-quantile, where for cost functions one would measure the risk through the \((1 - \alpha)\)-quantile.
We illustrate the application of the above results to robustness analysis with the help of the following examples.

**Example 3.4.** Consider the two-way network in Example 3.2. Let $h$ denote the stationary throughput at node $i = 3$ and model $\tau_2$ as being random. Then, the throughput at node 3 under stalling as a function of a certain $\tau_2$ is given by

$$h(\tau_2) = \eta_3 \frac{\rho_2}{\rho_2 + \tau_2},$$

and by computation

$$h^{-1}(y) = \frac{\eta_3 \rho_2}{y} - \rho_2 \quad \text{and} \quad \frac{d}{dy} h^{-1}(y) = -\frac{\eta_3 \rho_2}{y^2}.$$

In the following, two distributions for $\tau_2$ are elaborated, the uniform and exponential distribution, respectively:

- Assume that $\tau_2$ is uniformly distributed on $[a, b]$, with $0 < a < b < \infty$. Then,

  $$g(y) = \frac{1}{b-a} \frac{\eta_3 \rho_2}{y^2}, \quad \text{for} \quad \frac{\eta_3 \rho_2}{\rho_2 + b} \leq y \leq \frac{\eta_3 \rho_2}{\rho_2 + a},$$

  and zero otherwise. The value at risk, i.e., the $\alpha$-quantile of the stationary throughput, is also easily computable to be

  $$\text{VaR}(\alpha) = \frac{\eta_3 \rho_2}{\rho_2 + b - (b-a)\alpha},$$

  for $\alpha \in [0,1]$. In words, in $\alpha \cdot 100\%$ of the cases the actual throughput of the system will fall below $\text{VaR}(\alpha)$. Observe, that for $\alpha = 1$ we have $\text{VaR}(\alpha) = \eta_3 \rho_2 / (\rho_2 + a)$, which is the right bound of the support of $\tau_2$, and for $\alpha = 0$ we have $\text{VaR}(\alpha) = \eta_3 \rho_2 / (\rho_2 + b)$, which is the left bound of the support of $\tau_2$.

  Suppose $b = k \cdot a$ with $k > 1$ and we are rather uncertain about the true value of $\tau_2$, i.e., a relative large value for $k$ comes into question. Then the above analysis uncovers the exposed risk by expecting the throughput to be of order $h(\mu_{\tau_2})$ without taking the stochasticity into account. In particular, with chance $\alpha$ the realized throughput $h(\tau_2)$ is at least

  $$\left(1 - \frac{2\rho_2}{a} + k + \frac{1}{2} \frac{2\rho_2}{a} + 2(1-\alpha)k + 2\alpha\right) \cdot 100\%$$

  smaller than $h(\mu_{\tau_2})$. For example, let $\alpha = 0.1$, then with 0.1 probability the actual throughput $h(\tau_2)$ is at least approximately 44.4% smaller than $h(\mu_{\tau_2})$ for large $k$.

- Let $\tau_2$ be exponentially-$\lambda$-distributed. Then the density for the throughput via (3.18) equals

  $$g(y) = \frac{\lambda \eta_3 \rho_2}{y^2} \exp\left(-\lambda \rho_2 \left(\frac{\eta_3}{y} - 1\right)\right),$$

  for $y \in (0, \eta_3]$. It can be shown that the c.d.f. of $g(y)$ is given by

  $$G(y) = \exp\left(-\frac{\lambda \eta_3 \rho_2}{y} + \lambda \rho_2\right), \quad \text{for} \quad y \in (0, \eta_3],$$
which leads to
\[
\text{VaR}(\alpha) = \frac{\lambda \rho_2 \eta_3}{\lambda \rho_2 - \ln(\alpha)},
\]
for \(\alpha \in (0, 1)\).

Example 3.5. Consider the star network in Example 3.3. Let \(h\) denote the stationary throughput at node \(i = 2\) and model \(\tau_2\) as being random. Then, the throughput at node 2 as a function of certain \(\tau_2\) under skipping is
\[
h(\tau_2) = \eta_2 \pi(\emptyset) \left(1 + \frac{\tau_3}{\rho_3} + \frac{\tau_4}{\rho_4} + \frac{\tau_3 + \tau_4}{2 \min(\rho_3, \rho_4)}\right),
\]
with \(\pi(\emptyset)\) as given in (3.12), see Example 3.3.

Letting
\[
a_1 = 1 + \frac{\tau_3}{\rho_3} + \frac{\tau_4}{\rho_4} + \frac{\tau_3 + \tau_4}{2 \min(\rho_3, \rho_4)} + \frac{\tau_4}{2 \min(\rho_2, \rho_4)} + \frac{\tau_3}{2 \min(\rho_2, \rho_3)} + \frac{\tau_3 + \tau_4}{3 \min(\rho_2, \rho_3, \rho_4)},
\]
\[
a_2 = \frac{1}{\rho_2} + \frac{1}{2 \min(\rho_2, \rho_1)} + \frac{1}{2 \min(\rho_2, \rho_3)} + \frac{1}{3 \min(\rho_2, \rho_3, \rho_4)},
\]
and
\[
a_3 = \eta_2 \left(1 + \frac{\tau_3}{\rho_3} + \frac{\tau_4}{\rho_4} + \frac{\tau_3 + \tau_4}{2 \min(\rho_3, \rho_4)}\right),
\]
we may write for a certain \(\tau_2\)
\[
h(\tau_2) = \frac{a_3}{a_1 + a_2 \tau_2}.
\]
Hence,
\[
h^{-1}(y) = \frac{a_3}{a_2 y} - \frac{a_1}{a_2} \quad \text{and} \quad \frac{d}{dy} h^{-1}(y) = -\frac{a_3}{a_2 y^2}.
\]

In the following again uniformly and exponentially distributed \(\tau_2\) are considered, respectively:

- Assume that \(\tau_2\) is uniformly distributed on \([a, b]\), with \(0 < a < b < \infty\). Then,
\[
g(y) = \frac{1}{b - a} \frac{a_3}{a_2 y^2}, \quad \text{for} \quad \frac{a_3}{a_1 + a_2 b} \leq y \leq \frac{a_3}{a_1 + a_2 a}
\]
and zero otherwise. The value at risk, i.e., the \(\alpha\) quantile of the stationary throughput, is also easily computable to be
\[
\text{VaR}(\alpha) = \frac{a_3}{a_1 + a_2 b - \alpha(b - a)a_2},
\]
for \(\alpha \in [0, 1]\).

- For \(\tau_2\) exponentially distributed with parameter \(\lambda\) it holds that
\[
g(y) = \lambda \exp \left(-\lambda \left(\frac{a_3}{a_2 y} - \frac{a_1}{a_2}\right)\right) \frac{a_3}{a_2 y^2},
\]
for \( y \in (0, \frac{a_1}{a_2}) \), so that
\[
G(y) = \exp \left( -\lambda \left( \frac{a_3}{a_2 y} - \frac{a_1}{a_2} \right) \right),
\]
and thus
\[
\text{VaR}(\alpha) = \frac{\lambda a_3}{\lambda a_1 - a_2 \ln(\alpha)},
\]
for \( \alpha \in (0, 1) \).

For uniform and exponential distributions, \( \text{VaR}(\alpha) \) can be explicitly solved, which is due to the simplicity of both distributions. In the following we study the more challenging problem when the distribution of \( \tau_2 \) is assumed to be the truncated normal distribution. Revisit the two-way network from Example 3.4. Let \( \tau_2 \) be normally distributed with mean \( \mu \) and standard deviation \( \sigma \) but conditioned on interval \([\gamma_l, \gamma_r]\) where \( 0 \leq \gamma_l < \gamma_r \). The rational behind the conditioning is that negative values as well as non-realistically large values for \( \tau_2 \) are avoided. Then, following (3.18), the throughput at node 3 under stalling has density \( g \) for
\[
\frac{\eta_3 \rho_2}{\rho_2 + \gamma_r} \leq y \leq \frac{\eta_3 \rho_2}{\rho_2 + \gamma_l}
\]
given by
\[
g(y) = \frac{\eta_3 \rho_2}{\Delta \Phi \sigma \sqrt{2\pi} y^2} \exp \left( -\left( \frac{\eta_3 \rho_2 - \rho_2 - \mu}{2\sigma^2} \right)^2 \right), \tag{3.19}
\]
where
\[
\Delta \Phi = \Phi \left( \frac{\gamma_r - \mu}{\sigma} \right) - \Phi \left( \frac{\gamma_l - \mu}{\sigma} \right)
\]
and \( \Phi(\cdot) \) is the standard normal cumulative distribution function. To obtain the VaR we need to find the inverse of the cumulative distribution of the throughput given by
\[
G(y) = \int_{\frac{\eta_3 \rho_2}{\rho_2 + \gamma_r}}^{y} g(t) dt.
\]
Computing the inverse of a general function can usually only be performed numerically. However, in case the function of interest is analytical and can thus be written as a power series, a power series representation of the inverse can be obtained. This result is well-known in analysis, see, e.g., [64]. However, computing the actual elements of the power series is a challenging task. A first result can be found in Whittaker’s pioneering paper [190]. In particular, Whittaker provided an explicit expression for the elements of the power series of the inverse in terms of the elements of the power series of the original function. Unfortunately, the computation of the elements is rather demanding. An alternative approach, that suffers from the same computational burden is Lagrange’s inversion formula [3]. Dominici [65] introduced recently a new method for numerical inversion that is very well suited to computing the VaR of a random variable the density of which is of exponential form. In the following we will present this approach.
For an infinitely often differentiable mapping $f$ define the nested derivative $D^n[f](x)$ by the recursion

$$D^0[f](x) = 1$$

and

$$D^n[f](x) = \frac{d}{dx} \left( f(x)D^{n-1}[f](x) \right),$$

for $n \geq 1$. Let

$$h(x) = \int_a^x \frac{1}{f(t)} dt$$

with $f(a) \neq 0, \infty$. Then according to Theorem 4.1 in [65] the inverse of $h(x)$ is given by

$$h^{-1}(y) = a + f(a) \sum_{n \geq 1} D^{n-1}[f](a) \frac{y^n}{n!}, \quad (3.20)$$

where $|y| < \epsilon$ for some $\epsilon > 0$. The elements of the series can be easily evaluated by means of standard computer algebra tools. We refer to [65] for details.

**Example 3.6.** Consider the exponential mapping $e^x$ and apply the method of nested derivatives. Note that

$$h(x) := e^x - 1 = \int_0^x e^t dt.$$

Let $f(x) = e^{-x}$, then

$$D^n[f](x) = (-1)^n n! e^{-nx}$$

so that $D^n[f](0) = (-1)^n n!$. Since $f$ is analytical we obtain the inverse of $h(x)$ as

$$h^{-1}(y) = \sum_{n=1}^{\infty} (-1)^n \frac{y^n}{n},$$

which is easily recognizable as the series expansion of $\ln(y+1)$ around 0.

As illustrated in the above example, the method of nested derivatives allows for a direct analysis of the function under the integral. This is particularly useful in VaR computations as the analysis can directly be applied to the density and computation of the cumulative distribution function can thus be avoided.

In the following we present the main result on nested derivatives, where we write $\bar{g}(t) = 1/g(t)$.

**Theorem 3.5.** Let $G(y)$ be a cumulative distribution function on $B = [b_l, b_r]$, where $B = \mathbb{R}$ is not excluded. Suppose that there exits $g(t)$, for $t \in B$, such that

(i) for $y \in B$ it holds that

$$G(y) = \int_{b_l}^y g(t) dt,$$

(ii) $G$ is analytical on the interior of $B$ as a mapping in $y$,

(iii) there is an $a \in B$ such that $g(a) \neq 0$. 
Let 
\[ c_a = \int_{b}^{a} g(t) dt = G(a), \]
then
\[ \text{VaR}(\alpha) = a + \bar{g}(a) \sum_{n \geq 1} D^{n-1}[\bar{g}](a) \frac{(\alpha - c_a)^n}{n!}, \]
for \( \alpha \) sufficiently close to \( c_a \).

**Proof.** For \( x \geq a \), write
\[ G(x) = c_a + \int_{a}^{x} g(t) dt \]
and let
\[ G_{c_a}(x) = G(x) - c_a = \int_{a}^{x} \frac{1}{\bar{g}(t)} dt. \]
We now apply the nested derivatives method to \( \bar{g}(t) \). From (3.20) (see also [65]) it then follows that
\[ G_{c_a}^{-1}(y) = a + \bar{g}(a) \sum_{n \geq 1} D^{n-1}[\bar{g}](a) \frac{y^n}{n!}, \]
for \( |y| \) sufficiently small. Noting that \( \text{VaR}(\alpha) = G_{c_a}^{-1}(\alpha - c_a) \) concludes the proof. \( \square \)

Note that the advantage of Theorem 3.5 lies in the fact that the elements of the series expansion have to be computed once for \( a \), yielding a polynomial approximation of VaR on an entire interval. For ease of reference define for \( N \in \mathbb{N} \)
\[ \text{VaR}(N, \alpha) = a + \bar{g}(a) \sum_{n=1}^{N} D^{n-1}[\bar{g}](a) \frac{(\alpha - c_a)^n}{n!}, \]
so that \( \lim_{N \to \infty} \text{VaR}(N, \alpha) = \text{VaR}(\alpha) \).

**Example 3.7.** Reconsider the two-way network from Example 3.2, see also Example 3.4. Let \( \tau_2 \) be normally distributed with mean \( \mu \) and standard deviation \( \sigma \) but truncated on interval \( [\gamma_l, \gamma_r] \) where \( 0 \leq \gamma_l < \gamma_r \). See (3.19) for the density \( g(y) \) of the throughput. In order to approximate the VaR by Theorem 3.5, where \( a \) is chosen to be \( \frac{\eta_3 \rho_2}{\rho_2 + \gamma_r} \) so that \( c_a = 0 \) (note that this is allowed since \( g\left(\frac{\eta_3 \rho_2}{\rho_2 + \gamma_r}\right) \neq 0 \)), we will compute the series using the computer algebra algorithm provided in [65]. Using notation \( \bar{g}(y) = 1/g(y) \) it holds that
\[ \bar{g}(y) = \frac{\Delta \Phi}{\sigma \sqrt{2\pi}} \exp\left(\frac{\left(\frac{\eta_3 \rho_2}{\rho_2 + \gamma_r} - \rho_2 - \mu\right)^2}{2\sigma^2}\right), \]
for \( \frac{\eta_3 \rho_2}{\rho_2 + \gamma_r} \leq y \leq \frac{\eta_3 \rho_2}{\rho_2 + \gamma_r} \).
It follows from Maple calculations that \( \text{VaR}(1, \alpha) \), i.e., a VaR series approximation based on 1 term, equals
\[ \text{VaR}(1, \alpha) = a + \bar{g}(a) \alpha = \frac{\rho_2 \eta_3 \left(\Delta \Phi \sqrt{2\pi} \alpha \exp\left(\frac{(\gamma_r - \mu)^2}{2\sigma^2}\right) + \rho_2 + \gamma_r\right)}{(\rho_2 + \gamma_r)^2}. \]
For the VaR approximation of order 2 we have to add the term

$$
\bar{g}(a)D^{1}[\bar{g}](a) \frac{\alpha^2}{2} = \frac{-\eta_3 \rho_2 (\gamma_r^2 + (\rho_2 - \mu) \gamma_r - \rho_2 \mu - 2 \sigma^2) \alpha^2 \Delta_\Phi \pi \exp \left( \frac{(\gamma_r - \mu)^2}{2 \sigma^2} \right)}{2(\rho_2 + \gamma_r)^3}.
$$

In general, for the \( n \)-th term it holds

$$
\bar{g}(a)D^{n-1}[\bar{g}](a) \frac{\alpha^n}{n!} = \frac{(-1)^{n+1} \sigma^2 \eta_3 \rho_2 P(n)}{n!(\rho_2 + \gamma_r)} \left( \frac{\alpha \Delta_\Phi \sqrt{2 \pi} \exp \left( \frac{(\gamma_r - \mu)^2}{2 \sigma^2} \right)}{\sigma(\rho_2 + \gamma_r)} \right)^n,
$$

where \( P(n) \) is a homogeneous polynomial of degree \( 2(n-1) \) in variables \( \gamma_r, \sigma, \rho_2 \) and \( \mu \). In particular for \( P(n) \) with \( n = 1, 2, 3, 4 \) it holds

\[
\begin{align*}
P(1) &= 1 \\
P(2) &= \gamma_r^2 + (\rho_2 - \mu) \gamma_r - 2 \sigma^2 - \rho_2 \mu \\
P(3) &= 2 \gamma_r^4 + (4 \rho_2 - 4 \mu) \gamma_r^3 + (-5 \sigma^2 + 2 \rho_2^2 - 8 \rho_2 \mu + 2 \mu^2) \gamma_r^2 + (-4 \sigma^2 \rho_2 \\
&\quad+ 6 \sigma^2 \rho_2 - 4 \rho_2^2 \mu + 4 \rho_2 \mu^2) \gamma_r + 6 \sigma^4 + \sigma^2 \rho_2^2 \\
&\quad+ 6 \sigma^2 \rho_2 \mu + 2 \rho_2 \mu^2 \\
P(4) &= 6 \gamma_r^6 + (18 \rho_2 - 18 \mu) \gamma_r^5 + (-15 \sigma^2 + 18 \rho_2^2 - 54 \rho_2 \mu + 18 \mu^2) \gamma_r^4 + (-23 \sigma^2 \rho_2 \\
&\quad+ 37 \sigma^2 \mu + 6 \rho_2^3 + 54 \rho_2 \mu^2 + 54 \rho_2 \mu^2 - 6 \mu^3) \gamma_r^3 + (28 \sigma^4 - \sigma^2 \rho_2^2 \\
&\quad+ 67 \sigma^2 \rho_2 - 22 \sigma^2 \mu^2 - 18 \rho_2 \mu + 54 \rho_2 \mu^2 - 18 \rho_2 \mu^3) \gamma_r^2 \\
&\quad+ (20 \sigma^4 \rho_2 - 36 \sigma^4 \mu + 7 \sigma^2 \rho_2^3 + 23 \sigma^2 \rho_2 \mu - 44 \sigma^2 \rho_2 \mu^2 \\
&\quad+ 18 \rho_2 \mu^2 - 18 \rho_2 \mu^3) \gamma_r - 24 \sigma^4 \rho_2^2 - 8 \sigma^4 \rho_2 - 36 \sigma^4 \rho_2 \mu - 7 \sigma^2 \rho_2 \mu^2 \\
&\quad- 22 \sigma^2 \rho_2 \mu^2 - 6 \rho_2 \mu^3.
\end{align*}
\]

Figure 3.4 provides a numerical example which illustrates that the VaR series with a few terms already yields an accurate approximation for VaR(\( \alpha \)) with \( \alpha \in (0, 0.1) \). Specifically, we let \( \rho_2 = 1.1, \eta_3 = 1.5, \mu = 1, \sigma = 1.5 \) and \( [\gamma_l, \gamma_r] = [0.1, 2.75] \) so that \( \mu_{\gamma_2} = E[\tau_2] = 1.3259 \) and VaR(\( \gamma_2 \)) = 0.52134. Furthermore, the example illustrates that significant risk is ignored when taking \( h(\mu_{\gamma_2}) \) as measure for the throughput. Specifically, \( h(\mu_{\gamma_2}) \approx 0.68 \) whereas with probability 0.2 the actual throughput is approximately smaller than 0.52 (a difference of at least 23.5%) and with probability 0.1 the actual throughput is approximately smaller than 0.48 (a difference of at least 29.4%).

In case one is interested in VaR(\( \alpha \)) for \( \alpha \) around 0.3, Figure 3.4 shows that poor approximations are obtained via the series, even when using 10 terms for the series. The approximation for VaR(\( \alpha \)) with \( \alpha \) around 0.3 can be improved by choosing \( a \) in condition (iii) of Theorem 3.5 greater than \( b_l = \frac{n \rho_2}{\rho_2 + \gamma_r} \) such that \( c_a \) lies near 0.3. The downside is that this approach requires numerical evaluation of \( c_a \) and the search for an appropriate \( a \). But after this numerical burden, the series for VaR(\( \alpha \)) from Theorem 3.5 provides an accurate and efficient approximation for VaR(\( \alpha \)) with \( \alpha \) in a relative large interval around \( c_a \). As example, Figure 3.5 shows for the same instance as in Figure 3.4 that choosing \( a \) such that \( c_a \approx 0.3 \) leads to accurate approximations for VaR(\( \alpha \)) with \( \alpha \in (0.15, 0.45) \) even for a small number of series terms.
3.6 Conclusion

In this chapter we have integrated simultaneous breakdown and repair of servers in Jackson networks together with infinite supply servers in one framework. We obtained closed-form analytical solutions of the steady-state queue length distribution at stable
nodes. Numerical examples have illustrated that robustness analysis of such networks with respect to parameter insecurity is an insightful tool that becomes feasible due to the product-form type solution. Future research includes insecurity analysis of multiple parameters and further development of the risk analysis framework.