Master Thesis

Stochastic Performance Bounds for Independent Regulated Traffic in Queueing Networks:

Analytical Derivation, Parameter Optimization and Numerical Evaluation

Paul Nikolaus

December 2016

FACHBEREICH MATHEMATIK / INFORMATIK
TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

I hereby declare that this thesis entitled “Stochastic Performance Bounds for Independent Regulated Traffic in Queueing Networks: Analytical Derivation, Parameter Optimization, and Numerical Evaluation” is the result of my own research except as cited in the references. This thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :
Student : Paul Nikolaus
Student Number: 371932
Date : 13.12.2016

Supervisor : Prof. Dr. Ralf Korn

Co-Supervisor: Prof. Dr.-Ing. Jens B. Schmitt

Abstract

Network calculus (NC) is a young theory that was started in the 90s of the previous century. NC is a powerful theory to compute upper bounds on the performance of a queueing network based on the knowledge of so-called arrival and service curves. Today, there are two prevalent branches of NC: deterministic network calculus (DNC) and stochastic network calculus (SNC). DNC gives upper bounds that always hold whereas the SNC allows for violation for preset (usually low) probabilities.

In this thesis, we use mainly SNC to describe a system's behavior and to analyze its performance. It is divided into three main parts: Part I, "Introduction to Network Calculus", provides the reader with all the necessary definitions and results to grasp the rest of this thesis. Apart from this, we present the important mathematical tools we apply to analyze networks and give some examples where we illustrate how involved the calculation of the performance bounds can be.

The knowledge obtained in Part I is necessary for Part II, "Moment Generating Functions for Regulated Traffic", where we show that the bounds obtained in [MB00] and [VL01] can be transformed into MGF-bounds (this term is explained below). This enables us to implement their regulated traffic models quite easily since we only need the description of the MGF-bound.

Part III focuses on sink trees. We show how the different types of analyzes known from network calculus perform and that certain tendencies between the two prevailing approaches are clearly observable. In the second chapter of this part, we investigate in the field of parameter optimization in more detail and demonstrate that there is great potential for improvement by a thorough optimization.

In the last part, we show that the standard procedure, e.g., described in [Fid06], of computing the output bound, can be improved by utilizing the so-called Lyapunov inequality. We implemented the novel bound for different distributions (most of them are very common in network traffic modeling) and show that it can provide a significant improvement.

Acknowledgment

At first, I would like to thank my supervisors, Prof. Ralf Korn and Prof. Jens Schmitt who gave me the opportunity to do my research on various fields, but guided me through several imponderables. By this master thesis, I had the opportunity to experience new areas of academic work combined with newly added programming skills.

I am thankful to all the members of Prof. Schmitt's distributed computer systems (DISCO) lab in the computer science department, in particular Carolina Nogueira, Daniel Berger (who kindly shared his office with me), Hao Wang, Matthias Schäfer, Michael Beck, and, last but not least, Steffen Bondorf.

I would also like to thank the proofreaders of this thesis, David Osche, Prilly Oktoviany, and Thomas Ullmert who were an excellent help and improved the readability significantly.

My honest appreciation belongs to Álvaro Noreña, Jörg Wycisk, and Guido Knittel for the opportunity to do two internships during the master program.

Finally, I must express my very profound gratitude to my girlfriend, Yuki, and to my family, especially my mother, Irina, for always supporting me. The master's degree would not have been possible without you.

Paul Nikolaus, Kaiserslautern

Contents

List of Figures

Part I.

Introduction to Network Calculus

1. Background and Motivation

The origin of network calculus dates back to the (σ, ρ)-calculus presented in [Cru91a, Cru91b]. In these articles, the so-called "burstiness constraints" are introduced as well as tools to determine the delay of a network. Based on this, [PG93, PG94] combined general processor sharing (GPS) with leaky bucket admission control to make "worst-case performance guarantees" for networks. The usage of a service curve has been introduced there and generalized in [Cru95]. The analysis of min- and max-plus algebra has been examined in detail by the book of [BCOQ92]. As we see in this thesis, not only does the min-plus algebra simplify considerably the notation, it also allows to give a description of the network's service just by looking at it (of course, in case the topology is simple enough). [LBT96] and the book [LT01] provide an overview of deterministic network calculus' (DNC) state of the art at this point. Its probabilistic counterpart, stochastic network calculus (SNC), has been covered in the book by C.S. Chang, [Cha00], which also includes many DNC results, and more recently [JL08].

In deterministic network calculus, multiple approaches are known to obtain performance bounds of networks. The separate flow analysis (SFA) emerged from the need to "pay bursts only once" ([LT01]), i.e., that the burst part of the arrival curve is only accounted for once, since otherwise the analysis is much more conservative than the actual worst case. Preceding analyses like the total flow analysis (TFA) could not fulfill this property ([BS16]), on the other hand.

This could not prevent the SFA to struggle with overestimation even in tandems if multiplexing occurs, as was observed in [FS04, SZM08]. This issue can be tackled by the PMOO-SFA ("pay multiplexing only once"-SFA). We explain this type of analysis in detail in Chapter 2.

Unfortunately, for these most important types of algebraic network analyses, it has been shown that none of them dominates the other one ([SZF08]). Besides, the authors also resorted to an optimization approach. This idea has been followed up by replacing the well-known algebraic manipulations completely by optimization ([BJT10]). Unfortunately, the increased tightness of the bounds has been traded for a computational infeasibility for larger networks, as it has been demonstrated in [BNS16].

Initially, stochastic network calculus has been developed as a stochastic extension of network calculus ([Cha92a, Kur92, YS93]). Later on, it split into two branches: tail-bounds ([CBL05]) and MGF-bounds ([Fid06]). As it has been stated for SNC, "useful bounds demand the right approach and an optimal choice of parameters. Both steps are only partially known by now" ([Bec16]). This thesis is going to tackle the latter problem. Unfortunately, performance evaluation of general networks is not as evolved since dependencies are a highly disruptive factor. For example, for nonfeed-forward networks, additional constraints have to be imposed to guarantee the existence of a delay bound ([Cha92b, CLB00, LZ10]).

Of course, network calculus would not be considered as important as it is if it was not useful in any sort of application. Some examples are sensor network calculus (provides "the ability to derive deterministic statements" in sensor networks) [SR05], bandwidth estimation [LFV07], admission control [LWF96], switched Ethernets [SJH06], or, more recently, service level objectives in shared cloud networks [ZBHB16].

Application of Network Calculus does not only happen via "pen and paper" but is also executed through various software tools. There exists the RTC toolbox [WT06], the DISCO network calculator [SZ06, GZMS08]
[BS15], the COINC library [BT08], or the DISCO Stochastic Network Calculator [BS13], just to name a few.

Apart from the fundamental books [Cha00] and [LT01], several surveys give a good summary of NC's most important developments, such as [MP06, Fid10, FR15].

The fact that NC is trying to obtain bounds instead of describing the network properties exactly is an important contrast to the classical queueing theory ([Erl09, Erl17]) where, on the other hand, many more assumptions, e.g., on the arrival distribution, are necessary. This enables network calculus to do so in an elegant way:

1. It trades the accuracy of queueing theory for bounds, i.e., it relies on less assumptions but, therefore, only gives upper bounds instead of exact results.

2. The fact that bounds are at first computed for general arrival and services (distributional assumptions are used in a second step) provides a certain modularity and also an increased clarity.

2. SNC Background

This part is supposed to give a first introduction, in particular to the stochastic network calculus (SNC), since the entire thesis focuses on this branch of network calculus. We give the necessary definitions and results to provide the reader with a basic overview. Later, we introduce some core techniques to obtain performance bounds, the main target of this thesis.

The network topology is defined by a set of servers and their interactions with flows.

In this thesis, we only work on a discrete time set. We write $\mathbb{N}$ for the natural numbers without 0 ($\mathbb{N} = \{1, 2, \dots\}$) and $\mathbb{N}_0$ if 0 is included ($\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$). Similarly, we write $\mathbb{R}_0^+$ for $[0, \infty)$.

2.1. Arrivals Processes

Definition 2.1 (Arrival Flow)**.** Let $(a(i))_{i \in \mathbb{N}}$ be a sequence of non-negative real random variables, the so-called *increments*. We define a *flow* by the stochastic process A with time space $\mathbb{N}_0$ and state space $\mathbb{R}_0^+$ as

$$A(t) \coloneqq \sum_{i=1}^{t} a(i). \tag{2.1}$$

The bivariate version is defined as

$$A(s,t) \coloneqq A(t) - A(s) = \sum_{i=s+1}^{t} a(i).$$

In the following, we want to give the definition of a tail-bounded arrival process. It requires envelope and error functions.

Definition 2.2 (Tail-Bound)**.** A flow A is *tail-bounded* by an envelope $\alpha : \mathbb{N}_0 \times \mathbb{R}^+ \to \mathbb{R}_0^+$ with error $\eta : \mathbb{N}_0 \times \mathbb{R}^+ \to [0, 1]$, that is decreasing in both arguments, if for all $\varepsilon > 0$ and $s \leq t$ it holds that

$$\mathrm{P}(A(s,t) > \alpha(t-s, \varepsilon)) \leq \eta(t-s, \varepsilon).$$

Example 2.3. Assume a flow A with i.i.d. increments such that each of the increments' variance is bounded by $\sigma^2 < \infty$. Then we get a tail-bound by applying Chebyshev's inequality with $\alpha(t-s,\varepsilon) = (t-s)\cdot \mathrm{E}[a(1)] + \varepsilon$ and $\eta(t-s,\varepsilon) = \frac{1}{\varepsilon^2(t-s)}\sigma^2$:

$$
\begin{aligned}
\mathrm{P}(A(s,t) > (t-s)\left(\mathrm{E}[a(1)]+\varepsilon\right)) &\leq \mathrm{P}\left(\left|\frac{A(s,t)}{t-s} - \mathrm{E}[a(1)]\right| > \varepsilon\right) \\
&\overset{\text{(i.i.d.)}}{=} \mathrm{P}\left(\left|\frac{\sum_{i=s+1}^{t} a(i)}{t-s} - \mathrm{E}\left[\frac{\sum_{i=s+1}^{t} a(i)}{t-s}\right]\right| > \varepsilon\right) \\
&\leq \frac{1}{\varepsilon^2}\,\mathrm{var}\left(\frac{\sum_{i=s+1}^{t} a(i)}{t-s}\right) \\
&= \frac{1}{\varepsilon^2(t-s)^2}\,\mathrm{var}\left(\sum_{i=s+1}^{t} a(i)\right) \\
&\overset{\text{(i.i.d.)}}{=} \frac{1}{\varepsilon^2(t-s)^2}\sum_{i=s+1}^{t} \mathrm{var}(a(1)) \\
&= \frac{1}{\varepsilon^2(t-s)}\,\mathrm{var}(a(1)) \\
&\leq \frac{1}{\varepsilon^2(t-s)}\sigma^2.
\end{aligned}
$$

Now, we have seen one major brand of SNC, the tail-bound. A different approach would be to bound the arrivals' Moment Generating Function (MGF). In the following, we explain this approach as well as relation of tail-bounds and MGF-bounds. Afterwards, we also give their raisons d'être by pointing out their characteristic advantages and disadvantages.

Theorem 2.4 (Chernoff Bound and Moment Generating Functions)**.** *Let X be a real random variable and $a \in \mathbb{R}$, then the so-called Chernoff bound holds for all $\theta > 0$:*

$$\mathrm{P}(X > a) \leq e^{-\theta a}\phi_X(\theta), \tag{2.2}$$

where

$$\phi_X(\theta) \coloneqq \mathrm{E}\left[e^{\theta X}\right]$$

denotes the Moment Generating Function (MGF).

Proof. See [Ros06]. □

Remark 2.5. Due to the fact that (2.2) holds for an arbitrary parameter $\theta > 0$, we can reformulate it to

$$\mathrm{P}(X > a) \leq \inf_{\theta > 0}\left\{e^{-\theta a}\phi_X(\theta)\right\}. \tag{2.3}$$

Example 2.6. For

$$\mathrm{P}(A(s,t) > \alpha(t-s,\varepsilon)),$$

Inequality (2.2) leads to

$$\mathrm{P}(A(s,t) > \alpha(t-s,\varepsilon)) \leq e^{-\theta\alpha(t-s,\varepsilon)}\phi_{A(s,t)}(\theta).$$

Owing to the fact that the first factor decays exponentially, we concluded that in order to bound the probability, we can focus on bounding the MGF.

Definition 2.7 ($f(\theta, \cdot)$-Bound)**.** A flow A is $f(\theta, \cdot)$-*bounded* for $\theta > 0$, if for all $0 \leq s \leq t$ the MGF $\phi_{A(t-s)}(\theta)$ exists and

$$\phi_{A(s,t)}(\theta) \leq f(\theta, t-s)$$

holds.

Example 2.8. Let the increments $a(i)$, $1, \ldots, n$ be i.i.d. Bernoulli distributed with $\mathrm{P}(a(i) = 1) = p = 1 - \mathrm{P}(a(i) = 0)$. Then the MGF $\phi_{A(s,t)}(\theta)$ is $f(\theta, \cdot)$-bounded with $f(\theta, t-s) = \left(1 - p + pe^{\theta}\right)^{t-s}$.

This, as well as Example 2.6, shows the advantage of MGF-calculus: The fact that we can rely on the analysis of moment generating functions is a quite elegant way and additionally, there already exist many results about MGFs in mathematics that can be used.

Also, these two approaches, tail-bounds and MGF-bounds, do not just exist in parallel, they can be “transformed” into each other by the following theorem.

Theorem 2.9 (Conversion Theorem)**.** *If a flow A is tail-bounded by envelope α with error $\eta(t-s, \varepsilon) = \varepsilon$, it is also $f(\theta, t-s)$-bounded with*

$$f(\theta, t-s) = \int_0^1 e^{\theta \alpha(t-s, \varepsilon)} \mathrm{d}\varepsilon. \tag{2.4}$$

Conversely, if A is $f(\theta, t-s)$-bounded, it is also tail-bounded with $\alpha(t-s, \varepsilon) = \varepsilon$ and $\eta(t-s, \varepsilon) = f(\theta, t-s)e^{-\theta\varepsilon}$.

Proof. See [LBL07, Bec16]. □

So far, we have explained why MGF-bounds can give us bounds and we saw that they are expressed in an easier way than tail-bounds. We also showed how to convert MGF-bounds into tail-bounds without mentioning why tail-bounds can be advantageous in the first place. Therefore, take a look at the following example:

Example 2.10. Let the arrivals be an aggregation of two flows, i.e.,

$$A(s,t) = A_1(s,t) + A_2(s,t).$$

Obtaining a tail-bound, we look for an envelope α and error η s.t.

$$\mathrm{P}(A_1(s,t) + A_2(s,t) > \alpha(t-s, \varepsilon)) \leq \eta(t-s, \varepsilon).$$

An MGF-bound, on the other hand, is a function f s.t.

$$\mathrm{E}\left[e^{\theta(A_1(s,t) + A_2(s,t))}\right] \leq f(\theta, t-s).$$

This example, though very simple, is sufficient to explain some of the characteristics of tail-bounds and MGF-bounds. If the the "subflows" A_1 and A_2 are stochastically independent, then we can split

$$\mathrm{E}\left[e^{\theta(A_1(s,t)+A_2(s,t))}\right] = \mathrm{E}\left[e^{\theta A_1(s,t)}\right] \mathrm{E}\left[e^{\theta A_2(s,t)}\right]$$

and use, if known, the the subflows' MGFs which can sometimes easily lead to the desired bounds. But if this independence cannot be assumed, computation can become quite cumbersome needing Hölder's inequality, as we see below.

The tail-bound, on the other hand, requires a more involved computation needing envelope and error function that can sometimes be difficult to find. But in contrast to the MGF-calculus, no independence assumption on aggregated flows is necessary.

2.2. Service Processes

Summarizing the previous definitions and theorems, we have all the necessary tools to focus on servers and so-called service elements.

Definition 2.11 (Service Curve)**.** A service element offers a *service curve* $\beta : \mathbb{N}_0 \to \mathbb{R}_0^+$ if for all $t \in \mathbb{N}_0$ and all arrival flows A it holds true that:

$$A'(t) \geq \min_{0 \leq i \leq t} \left\{ A(i) + \beta(t-i) \right\}, \tag{2.5}$$

where A' denotes the output flow.

Example 2.12 (Rate-Latency Server)**.** A classic example is the rate-latency service curve

$$\beta_{R,T}(t) = R \cdot (t-T)^+$$

with $R, T \geq 0$. This means that at time t, the server waits T time units (latency) before providing a constant service with rate R. See therefore Figure 2.1.

We assume throughout this thesis that a server is "lossless", which means that no packets are lost on their way through the network. Everything that cannot be processed immediately will be queued and no traffic is created inside the system, i.e.,

$$A'(t) \leq A(t).$$

Definition 2.13 (Backlog and Delay)**.** Let $A(0,t)$ and $A'(0,t)$ be the arrival and departure process of a lossless server, respectively. The *backlog* at time $t \geq 0$ is

$$q(t) \coloneqq A(0,t) - A'(0,t).$$

The *virtual delay* at time $t \geq 0$ is defined as

$$d(t) \coloneqq \inf \left\{ s \geq 0 : A(0,t) \leq A'(0,t+s) \right\}.$$

If a first-in first-out (FIFO) ordering can be assumed, then we simply call d the *delay*.

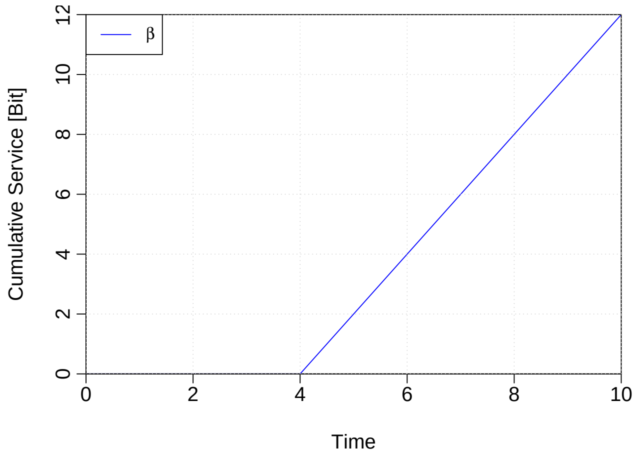


Figure 2.1.: Service curve β of a rate-latency server with $T = 4$ and $R = 2$.

Network Calculus shows its advantages when it comes to the analysis of the interaction between flows and a topology of servers, i.e., networks. Therefore, it uses operations called (de-)convolution and subtraction. We define them in the following and show why these operations are useful below.

Definition 2.14 (Bivariate Convolution in conventional Algebra)**.** The (de-)convolution of real-valued, bivariate functions $x(s,t)$ and $y(s,t)$ is defined as

$$\begin{aligned} x * y\,(s,t) &= \sum_{i=s}^{t} x(s,i) \cdot y(i,t) \\ x \circ y\,(s,t) &= \sum_{i=0}^{s} x(i,t) \cdot y(i,s) \end{aligned} \tag{2.6}$$

for $0 \leq s \leq t$.

Definition 2.15 (Bivariate Convolution in min-plus Algebra)**.** The min-plus (de-)convolution of real-valued, bivariate functions $x(s,t)$ and $y(s,t)$ is defined as

$$\begin{aligned} x \otimes y\,(s,t) &= \min_{s \leq i \leq t} \{x\,(s,i) + y\,(i,t)\} \\ x \oslash y\,(s,t) &= \max_{0 \leq i \leq s} \{x\,(i,t) - y\,(i,s)\} \end{aligned} \tag{2.7}$$

for $0 \leq s \leq t$.

The following proposition explains the relation between (2.6) and (2.7):

Proposition 2.16. *Let the random processes X, Y be independent. Then we get:*

$$
\begin{aligned}
&\phi_{X \otimes Y\,(s,t)}(-\theta) = \\
&\mathrm{E}\Big[e^{-\theta \min_{s \le i \le t}\{X(s,i)+Y(i,t)\}}\Big] \le \sum_{i=s}^{t} \mathrm{E}\Big[e^{-\theta X(s,i)}\Big] \cdot \mathrm{E}\Big[e^{-\theta Y(i,t)}\Big] \\
&\qquad = \left(\phi_X(-\theta) * \phi_Y(-\theta)\right)(s,t)
\end{aligned}
$$

for all $\theta > 0$ such that the above MGFs exists. Further, we obtain:

$$
\begin{aligned}
&\phi_{X \oslash Y\,(s,t)}(\theta) = \\
&\mathrm{E}\Big[e^{\theta \max_{0 \le i \le s}\{X(i,t)-Y(i,s)\}}\Big] \le \sum_{i=0}^{s} \mathrm{E}\Big[e^{\theta X(i,t)}\Big] \cdot \mathrm{E}\Big[e^{-\theta Y(i,s)}\Big] \\
&\qquad = \left(\phi_X(\theta) \circ \phi_Y(-\theta)\right)(s,t)
\end{aligned}
$$

for all $\theta > 0$ such that the above MGFs exists.

Proof. Lemma 1 in [Fid06]. □

Additionally, we need a bivariate version of (2.5):

Definition 2.17 (Dynamic S-Server)**.** Assume a service element has an arrival flow A as its input and the output is denoted by A'. Let S be a stochastic process with:

$$S(s,t) \le S\,(s,u) \quad \forall t \le u \in \mathbb{N}_0 \quad a.s.$$

The service element is a *dynamic S-server* if for all $t \ge 0$ it holds that:

$$A'(0,t) \ge A \otimes S\,(0,t) \overset{(2.7)}{=} \min_{0 \le i \le t} \{A\,(0,i) + S\,(i,t)\}\,. \tag{2.8}$$

We also assume that a server is non-idling, i.e., if any arrivals are at the server to be processed, the server is busy with this task.

Example 2.18 (Constant Rate Service)**.** Probably the easiest example of a dynamic S-server is a server providing constant service, i.e.,

$$S(s,t) = r \cdot (t-s)$$

for a constant $r \ge 0$. This service rate is assumed throughout this whole thesis.

Example 2.19 (Prioritized Cross Traffic)**.** Let us now take a look at Figure 2.2. There, we have two input flows, f_2, and the so-called *flow of interest* (foi), f_1. Let us assume that the arrival of flows f_2 is prioritized such that the S_1 serves all A_2 before A_1. Additionally, the service element, if busy, provides a constant service. How can the service for flow f_1 be described? One can conclude that the service element is a dynamic S-server for A_1 with

$$S(s,t) \coloneqq \left[c \cdot (t-s) - A_2(s,t)\right]^+ .$$

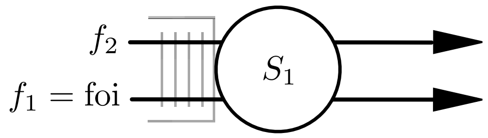


Figure 2.2.: One server with two input flows.

Proof. See Proposition 1 in [Fid06]. □

Similar to arrival flows' tail-bounds and MGF-bounds (Def. 2.2 and Def. 2.7), there exists a version for dynamic S-servers:

Definition 2.20 (Tail-Bounds for Service Curves)**.** A dynamic S-server is tail-bounded by envelope α with error η, if for all arrival flows A and all $t \in \mathbb{N}_0$, $\varepsilon > 0$ it holds that:

$$\mathrm{P}\Big(A'(0,t) < A \otimes (S - \alpha(\varepsilon))\ (0,t)\Big) < \eta\,(t,\varepsilon)\,.$$

Definition 2.21 ($f(\theta,\cdot)$-Bound for Service Curves)**.** A dynamic S-server is $f(\theta,t)$-bounded for some $\theta > 0$, if $\phi_{S(s,t)}(-\theta)$ exists and

$$\phi_{S(s,t)}(-\theta) \leq f(\theta, t-s)$$

for all $0 \leq s \leq t$.

With all these definitions, we are able to state the Concatenation Theorem as in [Fid06].

Theorem 2.22 (Concatenation Theorem)**.** *Let there be two service elements, such that the output of the first service element is the input for the second service element (as in Figure 2.3). Assume the first element to be a dynamic S^1-server and the second to be a dynamic S^2-server. Then the whole system results in a dynamic $S^1 \otimes S^2$-server.*

Proof. See Theorem 1 in [Fid06]. □

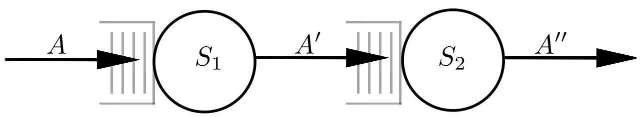


Figure 2.3.: Setting in Concatenation Theorem.

2.3. Performance Bounds

The following theorem is a fundamental result in network calculus. It tells us how to obtain performance bounds by just having knowledge about the arrivals and the service and also shows why network calculus is considered to be a very "elegant" framework.

Theorem 2.23 (Backlog, Delay, and Output Bounds)**.** *Consider a dynamic S-server $S(s,t)$ with arrival process $A(s,t)$. The backlog at time $t \geq 0$ is upper bounded by*

$$q(t) \leq A \oslash S\,(t,t). \tag{2.9}$$

If we assume FIFO scheduling, then the delay at $t \geq 0$ is upper bounded by

$$d(t) \leq \inf\left\{s \geq 0 : A \oslash S\,(t+s,t) \leq 0\right\}. \tag{2.10}$$

The departure process A' is upper bounded for any $0 \leq s \leq t$ according to

$$A'(s,t) \leq A \oslash S\,(t,t). \tag{2.11}$$

Proof. See Theorem 2 in [Fid06]. □

Remark 2.24. Theorem 2.23 already reveals the lack of tightness as probably the main flaw of network calculus. Let us for example take a look at the output bound (2.11):

$$A'(s,t) \leq A \oslash S\,(s,t) = \max_{0 \leq i \leq s}\left\{A\,(i,t) - S\,(i,s)\right\}.$$

Since $i = s$ is included in this set (that is maximized), we already know that

$$\max_{0 \leq i \leq s}\left\{A\,(i,t) - S\,(i,s)\right\} \geq A(s,t) - S\,(s,s) = A(s,t).$$

This means that our output bound is at least as big as the input. Knowing that no data is created inside network, i.e.,

$$A(s,t) \geq A'(s,t),$$

we obtain

$$A'(s,t) \leq A(s,t) \leq \max_{0 \leq i \leq s}\left\{A(i,t) - S(i,s)\right\} = A \oslash S\,(s,t).$$

This already indicates how conservative the bounds can be, based on the observation that the output bound is always greater than or equal to the input.

Since we only consider MGF-bounds in this thesis, we need an MGF - version of Theorem 2.23:

Theorem 2.25 (MGF-Bounds for Backlog, Delay, and Output)**.** *Consider a dynamic S-server $S(s,t)$ with arrival process $A(s,t)$. The backlog at time $t \geq 0$ is upper bounded by*

$$\mathrm{P}(q(t) > x) \leq e^{-\theta x}\,\mathrm{E}\Big[e^{\theta A \oslash S\,(t,t)}\Big]. \tag{2.12}$$

If we assume FIFO scheduling, then the delay at $t \geq 0$ is upper bounded by

$$\mathrm{P}(d(t) > T) \leq \mathrm{E}\Big[e^{\theta A \oslash S\,(t+T,t)}\Big]. \tag{2.13}$$

The departure process A' is upper bounded for any $0 \leq s \leq t$ according to

$$\phi_{A'(s,t)}(\theta) \leq \mathrm{E}\Big[e^{\theta A \oslash S\,(s,t)}\Big]. \tag{2.14}$$

Proof. The proof is given in [Fid06, Bec16]. For each bound, one starts with the result of Theorem 2.23 and continues with standard probability inequalities. □

Remark 2.26. As it is explained in Chapter 2.4, the convolution of multiple servers is a basic operation providing us with the possibility to obtain performance bounds. Since we are dealing in this thesis with MGF-bounds, the conversion from the max in the convolution to the sum is not obvious, as we see here:

$$\begin{aligned} A \otimes B \otimes C\,(s,t) &= \min_{s \le k_1 \le t} \{A\,(s,k_1) + B \otimes C\,(k_1,t)\} \\ &= \min_{s \le k_1 \le t} \left\{A\,(s,k_1) + \min_{k_1 \le k_2 \le t} \{B\,(k_1,k_2) + C\,(k_2,t)\}\right\} \\ &= \min_{s \le k_1 \le k_2 \le t} \{A\,(s,k_1) + B\,(k_1,k_2) + C\,(k_2,t)\} \end{aligned} \tag{2.15}$$

gives us

$$\begin{aligned} \mathrm{E}\Big[e^{-\theta(A\otimes B\otimes C\,(s,t))}\Big] &\overset{(2.15)}{=} \mathrm{E}\Big[e^{-\theta \min_{s\le k_1\le t}\left\{A(s,k_1)+\min_{k_1\le k_2\le t}\{B(k_1,k_2)+C(k_2,t)\}\right\}}\Big] \\ &= \mathrm{E}\Big[e^{\theta \max_{s\le k_1\le t}\left\{-A(s,k_1)-\min_{k_1\le k_2\le t}\{B(k_1,k_2)+C(k_2,t)\}\right\}}\Big] \\ &= \mathrm{E}\Big[e^{\theta \max_{s\le k_1\le t}\left\{-A(s,k_1)+\max_{k_1\le k_2\le t}\{-B(k_1,k_2)-C(k_2,t)\}\right\}}\Big] \\ &\le \sum_{s\le k_1\le t} \mathrm{E}\Big[e^{\theta\left(-A(s,k_1)+\max_{k_1\le k_2\le t}\{-B(k_1,k_2)-C(k_2,t)\}\right)}\Big] \\ &= \sum_{s\le k_1\le t}\sum_{k_1\le k_2\le t} \mathrm{E}\Big[e^{-\theta(A(s,k_1)+B(k_1,k_2)+C(k_2,t))}\Big] \\ &= \sum_{s\le k_1\le k_2\le t} \mathrm{E}\Big[e^{-\theta(A(s,k_1)+B(k_1,k_2)+C(k_2,t))}\Big] \end{aligned}$$

with

$$\sum_{k\le i\le j\le n} a_{i,j} := \sum_{i=k}^{n}\sum_{j=i}^{n} a_{i,j} = \sum_{j=k}^{n}\sum_{i=k}^{j} a_{i,j}. \tag{2.16}$$

Proposition 2.27 (n-fold Convolution Bound)**.** *For the n-fold convolution*

$$\begin{aligned} A_1\otimes\cdots\otimes A_n\,(s,t) &= \min_{s\le k_1\le t}\left\{A_1\,(s,k_1) + \ldots \min_{k_{n-2}\le k_{n-1}\le t}\{A_{n-1}\,(k_{n-2},k_{n-1}) + A_n\,(k_{n-1},t)\}\right\} \\ &= \min_{s\le k_1\le\cdots\le k_{n-1}\le t}\{A_1\,(s,k_1)+\cdots+A_{n-1}\,(k_{n-2},k_{n-1})+A_n\,(k_{n-1},t)\} \end{aligned}$$

we get

$$\begin{aligned} \mathrm{E}\Big[e^{-\theta(A_1\otimes\cdots\otimes A_n\,(s,t))}\Big] &\le \sum_{s\le k_1\le t}\cdots\sum_{k_{n-2}\le k_{n-1}\le t} \mathrm{E}\Big[e^{-\theta(A_1(s,k_1)+\cdots+A_n(k_{n-1},t))}\Big]. \\ &\le \sum_{s\le k_1\le\cdots\le k_{n-1}\le t} \mathrm{E}\Big[e^{-\theta(A_1(s,k_1)+\cdots+A_n(k_{n-1},t))}\Big]. \end{aligned} \tag{2.17}$$

Proof. We prove the proposition via induction. We start with $n = 2$, since we need at least two arrivals for the convolution operation. But this case is already proved in Proposition 2.16. Now we assume that (2.17) holds for an arbitrary but fixed $n \geq 2$ and perform the induction step:

$$
\begin{aligned}
\mathrm{E}\left[e^{-\theta(A_1 \otimes \cdots \otimes A_{n+1}\,(s,t))}\right] &\overset{\text{(IS)}}{\leq} \sum_{s \leq k_1 \leq t} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq t} \mathrm{E}\left[e^{-\theta(A_1(s,k_1)+\cdots+A_n \otimes A_{n+1}\,(k_{n-1},t))}\right] \\
&= \sum_{s \leq k_1 \leq t} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq t} \mathrm{E}\left[e^{-\theta\left(\min_{k_{n-1} \leq k_n \leq t}\{A_1(s,k_1)+\cdots+A_n(k_{n-1},k_n)+A_{n+1}(k_n,t)\}\right)}\right] \\
&= \sum_{s \leq k_1 \leq t} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq t} \mathrm{E}\left[e^{\theta\left(\max_{k_{n-1} \leq k_n \leq t}\{-A_1(s,k_1)-\cdots-A_n(k_{n-1},k_n)-A_{n+1}(k_n,t)\}\right)}\right] \\
&\leq \sum_{s \leq k_1 \leq t} \cdots \sum_{k_{n-1} \leq k_n \leq t} \mathrm{E}\left[e^{\theta(-A_1(s,k_1)-\cdots-A_{n+1}(k_n,t))}\right] \\
&= \sum_{s \leq k_1 \leq t} \cdots \sum_{k_{n-1} \leq k_n \leq t} \mathrm{E}\left[e^{-\theta(A_1(s,k_1)+\cdots+A_{n+1}(k_n,t))}\right].
\end{aligned}
$$

□

Remark 2.28. We would like to point out to the reader that the term network calculus is not to be confused with network theory or network optimization, even though some similarities might appear. For example problem such as the max flow show a certain resemblance, but concepts and techniques to tackle this problem are completely different.

2.4. Analyzing Larger Topologies

So far, we have mostly considered a simple one-server topology. Now, we are going to show how to obtain performance bounds in settings like the 2-chain tandem in Figure 2.4 and also look at more general topologies. We abbreviate the flow of interest by "foi".

The main method is to apply subtraction and convolution operations to the network that enable us to apply the techniques we know from the one-server setting. In a second step, we make frequent use of certain well-known inequalities we restate in Section 2.5.

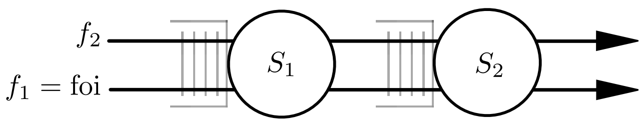


Figure 2.4.: 2-chain tandem.

2.4.1. Separated Flow Analysis (SFA)

The separated flow analysis (SFA) works, loosely speaking, according to the "subtract first, then convolute"-policy. This means that at each server, we subtract the flows crossing the foi and afterwards, we convolute all the servers along the foi. Applying this to the "2-chain tandem" in Figure 2.4, the leftover service curve for the foi is

$$S_{\text{l.o.}} = [S_1 - A_2]^+ \otimes [S_2 - (A_2 \oslash S_1)]^+ . \tag{2.18}$$

We can already see the basic mechanics of the SFA in this example: First, we subtract the arrivals of cross flow f_2, A_2, at the first server, and then $A_2 \oslash S_1$ at the second server. $A_2 \oslash S_1$ resembles the output of the prioritized cross traffic, as in (2.11). At this point, we would like to mention that the A_2-part appears twice. This usually leads to additional dependencies in the moment generating functions, as we see below.

2.4.2. Pay Multiplexing Only Once-SFA (PMOO-SFA)

The pay multiplexing only once - separated flow analysis (PMOO-SFA) tries the complete opposite compared to the SFA: It convolves (if possible) all the servers at first and subtracts the cross traffic at the very last step ([SZF08]). For the network depicted in Figure 2.4, the PMOO-SFA service curve has the following shape:

$$S_{\text{l.o.}} = [S_1 \otimes S_2 - A_2]^+ . \tag{2.19}$$

But convolving the servers at first is not that simple since it requires them to share the cross flow. If this is not the case, residual cross traffic has to be subtracted, usually leaving a lot of options to try.

2.5. Applied Inequalities

Even after simplifying a network to a one-server topology, e.g., by applying the SFA or the PMOO-SFA, we still end up in an MGF-Bound with a huge term we cannot quantify. Due to the fact that we make assumptions on our arrival and service distribution, we are able to evaluate terms of the form $\mathrm{E}\left[e^{\theta A(s,t)}\right]$ and $\mathrm{E}\left[e^{\theta S(s,t)}\right]$. As a result, we apply several well-known inequalities so we obtain bounds that consist of exactly these aforementioned terms, $\mathrm{E}\left[e^{\theta A(s,t)}\right]$ and $\mathrm{E}\left[e^{\theta S(s,t)}\right]$.

Theorem 2.29 (Hölder's Inequality)**.** *Let $p, q > 1$ such that $\frac{1}{q} + \frac{1}{p} = 1$ and assume the p-th moment of X and the q-th moment of Y to be finite. Then Hölder's inequality holds:*

$$\mathrm{E}[XY] \leq \mathrm{E}[|XY|] \leq \mathrm{E}[|X|^p]^{\frac{1}{p}} \, \mathrm{E}[|Y|^q]^{\frac{1}{q}} . \tag{2.20}$$

Proof. Theorem 23.10 in [JP02]. □

Proposition 2.30 (Quasi-Union Bound)**.** *For all $x_i \in \mathbb{R}$, $i = 1, \ldots, n$ and $\theta > 0$, it holds true that*

$$e^{\theta \max_{i=1,\ldots,n} x_i} \leq \sum_{i=1}^{n} e^{\theta x_i}. \tag{2.21}$$

For convenience, we call this the "quasi-union bound".

Proof. The terms $e^{\theta x_i}$ are greater than 0 for all $i = 1, \ldots, n$. This leads to

$$e^{\theta x_i} \leq \sum_{i=1}^{n} e^{\theta x_i} \qquad \forall i = 1, \ldots, n$$

and thus the maximum of the left hand side is smaller than the sum. Finally, we use

$$\max_{i=1,\ldots,n} \left\{ e^{\theta x_i} \right\} = e^{\theta \max_{i=1,\ldots,n}\{x_i\}}.$$

□

Remark 2.31. Let X_1, X_2 and X_3 be random variables such that the $p_1 \cdot p_2 \cdot p_3 \cdot p_4$-th moment exists ($p_i > 1$, $i = 1, \ldots, 4$). It is tempting to apply Hölder's inequality twice which gives us

$$\begin{aligned}
\mathrm{E}[|X_1| \cdot |X_2| \cdot |X_3|] &\leq (\mathrm{E}[|X_1|^{p_1}])^{\frac{1}{p_1}} \left(\mathrm{E}[(|X_2| \cdot |X_3|)^{p_2}]\right)^{\frac{1}{p_2}} \\
&\leq (\mathrm{E}[|X_1|^{p_1}])^{\frac{1}{p_1}} \left(\mathrm{E}[|X_2|^{p_2 p_3}]\right)^{\frac{1}{p_2 p_3}} \left(\mathrm{E}[|X_3|^{p_2 p_4}]\right)^{\frac{1}{p_2 p_4}}
\end{aligned}$$

with

$$\begin{aligned}
1 =& \frac{1}{p_1} + \frac{1}{p_2} \\
1 =& \frac{1}{p_3} + \frac{1}{p_4}.
\end{aligned}$$

But this raises several new questions, namely how to order the X_i. By commutativity of the product, the order does not affect the value on the left hand side, whereas the different powers p_i may lead to different results on the right hand side. As we see below, our remedy is a generalized version of Hölder's inequality.

Proposition 2.32 (Generalized Hölder Inequality)**.** *Let $X_1, \ldots, X_n$ such that $X_i \in L^{p_i}$ be random variables. Then we have*

$$\mathrm{E}\left[\prod_{i=1}^{n} |X_i|\right] \leq \prod_{i=1}^{n} \mathrm{E}\left[|X_i|^{p_i}\right]^{\frac{1}{p_i}} \tag{2.22}$$

with

$$\sum_{i=1}^{n} \frac{1}{p_i} = 1$$

and $p_i > 0$.

Proof. A proof is given in [Che01]. □

2.6. 2-Chain Tandem

Together with Theorem 2.25, the SFA (Section 2.4.1) as well as the PMOO-SFA (Section 2.4.2) allows for computing performance bounds in networks. In the "2-chain tandem" in Figure 2.4, this leads to bounds shown in the following. We assume the arrivals and servers to be independent.

Proposition 2.33 (Backlog Bound with SFA)**.** *For the backlog bound, we obtain*

$$\mathrm{P}(q(t) > x) \leq e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t} \left(\mathrm{E}\left[e^{p\theta A_2(k,i)}\right]^{\frac{1}{p}} \mathrm{E}\left[e^{-p\theta S_1(k,i)}\right]^{\frac{1}{p}} \right.\right.$$
$$\left.\left. \cdot \left(\sum_{j=0}^{i} \mathrm{E}\left[e^{q\theta A_2(j,t)}\right] \mathrm{E}\left[e^{-q\theta S_1(j,i)}\right] \right)^{\frac{1}{q}} \mathrm{E}\left[e^{-q\theta S_2(i,t)}\right]^{\frac{1}{q}} \right) \right),$$

with

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Proof. We compute

$$\begin{aligned}
\mathrm{P}(q(t) > x) &\overset{(2.12)}{\leq} e^{-\theta x}\, \mathrm{E}\left[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o.}}(t,t)}\right] \\
&\overset{(2.7)}{=} e^{-\theta x}\, \mathrm{E}\left[e^{\theta \max_{0\leq k\leq t}\{A_{\mathrm{foi}}(k,t) - S_{\mathrm{l.o.}}(k,t)\}}\right] \\
&\overset{(2.18)}{=} e^{-\theta x}\, \mathrm{E}\left[e^{\theta \max_{0\leq k\leq t}\left\{A_{\mathrm{foi}}(k,t) - [S_1 - A_2]^+ \otimes [S_2 - (A_2 \oslash S_1)]^+(k,t)\right\}}\right] \\
&\overset{(2.21)}{\leq} e^{-\theta x} \sum_{k=0}^{t} \mathrm{E}\left[e^{\theta\left(A_1(k,t) - \left([S_1 - A_2]^+ \otimes [S_2 - (A_2 \oslash S_1)]^+\,(k,t)\right)\right)}\right] \\
&\overset{\text{(indep.)}}{=} e^{-\theta x} \sum_{k=0}^{t} \mathrm{E}\left[e^{\theta A_1(k,t)}\right] \mathrm{E}\left[e^{-\theta\left([S_1 - A_2]^+ \otimes [S_2 - (A_2 \oslash S_1)]^+\,(k,t)\right)}\right] \\
&\overset{(2.21)}{\leq} e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t} \mathrm{E}\left[e^{-\theta\left([S_1(k,i) - A_2(k,i)]^+ + [S_2(i,t) - (A_2 \oslash S_1)(i,t)]^+\right)}\right] \right) \\
&= e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t} \mathrm{E}\left[e^{-\theta [S_1(k,i) - A_2(k,i)]^+} e^{-\theta [S_2(i,t) - A_2 \oslash S_1\,(i,t)]^+}\right] \right).
\end{aligned}$$

Here, we have stochastic dependence even though we assumed the arrivals and servers to be independent. It is a dependence that is caused by the SF-Analysis since it does not happen for the PMOO-SFA, as we see below. These dependent factors in the expectation are the reasons why we proceed by applying Hölder's inequality (2.20):

$$
\begin{aligned}
\ldots &\overset{(2.20)}{\leq} e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \sum_{i=k}^{t} \left(\mathrm{E}\Big[e^{-p\theta[S_1(k,i)-A_2(k,i)]^+}\Big]\right)^{\frac{1}{p}} \left(\mathrm{E}\Big[e^{-q\theta[S_2(i,t)-(A_2 \oslash S_1)(i,t)]^+}\Big]\right)^{\frac{1}{q}} \right) \\
&= e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \sum_{i=k}^{t} \left(E\left[e^{p\theta \min\{0, A_2(k,i)-S_1(k,i)\}}\right]\right)^{\frac{1}{p}} \left(E\left[e^{q\theta \min\{0,(A_2 \oslash S_1)(i,t)-S_2(i,t)\}}\right]\right)^{\frac{1}{q}} \right) \\
&\leq e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \sum_{i=k}^{t} \left(E\left[e^{p\theta (A_2(k,i)-S_1(k,i))}\right]\right)^{\frac{1}{p}} \left(E\left[e^{q\theta ((A_2 \oslash S_1)(i,t)-S_2(i,t))}\right]\right)^{\frac{1}{q}} \right) \\
&= e^{-\theta x} \sum_{k=0}^{n} \left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \sum_{i=k}^{t} \mathrm{E}\Big[e^{p\theta A_2(k,i)}\Big]^{\frac{1}{p}} \mathrm{E}\Big[e^{-p\theta S_1(k,i)}\Big]^{\frac{1}{p}} \mathrm{E}\Big[e^{q\theta (A_2 \oslash S_1)(i,t)}\Big]^{\frac{1}{q}} \mathrm{E}\Big[e^{-q\theta S_2(i,t)}\Big]^{\frac{1}{q}} \right) \\
&\overset{(2.21)}{\leq} e^{-\theta x} \sum_{k=0}^{t} \Bigg(E\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t} \Bigg(\mathrm{E}\Big[e^{p\theta A_2(k,i)}\Big]^{\frac{1}{p}} \mathrm{E}\Big[e^{-p\theta S_1(k,i)}\Big]^{\frac{1}{p}} \\
&\quad \left(\sum_{j=0}^{i} \mathrm{E}\Big[e^{q\theta(A_2(j,t)-S_1(j,i))}\Big]\right)^{\frac{1}{q}} \mathrm{E}\Big[e^{-q\theta S_2(i,t)}\Big]^{\frac{1}{q}} \Bigg)\Bigg) \\
&= e^{-\theta x} \sum_{k=0}^{t} \Bigg(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \sum_{i=k}^{t} \Bigg(\mathrm{E}\Big[e^{p\theta A_2(k,i)}\Big]^{\frac{1}{p}} \mathrm{E}\Big[e^{-p\theta S_1(k,i)}\Big]^{\frac{1}{p}} \\
&\quad \cdot \left(\sum_{j=0}^{i} \mathrm{E}\Big[e^{q\theta A_2(j,t)}\Big] \mathrm{E}\Big[e^{-q\theta S_1(j,i)}\Big]\right)^{\frac{1}{q}} \mathrm{E}\Big[e^{-q\theta S_2(i,t)}\Big]^{\frac{1}{q}} \Bigg)\Bigg),
\end{aligned}
$$

with

$$\frac{1}{p} + \frac{1}{q} = 1.$$

□

Proposition 2.34 (Backlog Bound with PMOO-SFA)**.** *The PMOO-SFA gives us*

$$\mathrm{P}(q(t) > x) \leq e^{-\theta x} \sum_{k=0}^{t} \left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \mathrm{E}\Big[e^{\theta A_2(k,t)}\Big] \sum_{i=k}^{t} \mathrm{E}\Big[e^{-\theta S_1(k,i)}\Big] \mathrm{E}\Big[e^{-\theta S_2(i,t)}\Big] \right).$$

Proof. We get

$$
\begin{aligned}
\mathrm{P}(q(t) > x) &\overset{(2.12)}{\leq} e^{-\theta x} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o.}}(t,t)}\Big] \\
&\overset{(2.7)}{=} e^{-\theta x} \mathrm{E}\Big[e^{\theta \max_{0 \leq k \leq t}\{A_{\mathrm{foi}}(k,t) - S_{\mathrm{l.o.}}(k,t)\}}\Big] \\
&\overset{(2.19)}{=} e^{-\theta x} \mathrm{E}\Big[e^{\theta \max_{0 \leq k \leq t}\left\{A_1(k,t) - [S_1 \otimes S_2\,(k,t) - A_2(k,t)]^+\right\}}\Big]
\end{aligned}
$$

$$
\begin{aligned}
&= e^{-\theta x}\,\mathrm{E}\Big[\max_{0\le k\le t}\Big\{e^{\theta\left(A_1(k,t)-[S_1\otimes S_2\,(k,t)-A_2(k,t)]^+\right)}\Big\}\Big]\\
&\overset{(2.21)}{\le} e^{-\theta x}\sum_{k=0}^{t}\mathrm{E}\Big[e^{\theta\left(A_1(k,t)-[S_1\otimes S_2\,(k,t)-A_2(k,t)]^+\right)}\Big]\\
&\overset{\text{(indep.)}}{=} e^{-\theta x}\sum_{k=0}^{t}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{-\theta[S_1\otimes S_2\,(k,t)-A_2(k,t)]^+}\Big]\\
&= e^{-\theta x}\sum_{k=0}^{t}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{-\theta\max\{0,S_1\otimes S_2\,(k,t)-A_2(k,t)\}}\Big]\\
&= e^{-\theta x}\sum_{k=0}^{t}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{\theta\min\{0,A_2(k,t)-S_1\otimes S_2\,(k,t)\}}\Big]\\
&\le e^{-\theta x}\sum_{k=0}^{t}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{\theta(A_2(k,t)-S_1\otimes S_2\,(k,t))}\Big]\\
&\overset{\text{(indep.)}}{=} e^{-\theta x}\sum_{k=0}^{t}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{\theta A_2(k,t)}\Big]\,\mathrm{E}\Big[e^{-\theta S_1\otimes S_2\,(k,t)}\Big]\\
&\overset{(2.21)}{\le} e^{-\theta x}\sum_{k=0}^{t}\left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{\theta A_2(k,t)}\Big]\sum_{i=k}^{t}\mathrm{E}\Big[e^{-\theta(S_1(k,i)+S_2(i,t))}\Big]\right)\\
&\overset{\text{(indep.)}}{=} e^{-\theta x}\sum_{k=0}^{t}\left(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big]\,\mathrm{E}\Big[e^{\theta A_2(k,t)}\Big]\sum_{i=k}^{t}\mathrm{E}\Big[e^{-\theta S_1(k,i)}\Big]\,\mathrm{E}\Big[e^{-\theta S_2(i,t)}\Big]\right).
\end{aligned}
$$

□

Remark 2.35. We can already see an important difference to the SFA: The fact that we convoluted first in the PMOO-SFA spared us the dependency that occurred with the SFA. This is an important factor in our SFA / PMOO-SFA comparison later on. In DNC, it has been proven that neither the SFA nor the PMOO-SFA is always better than the other [SZF08]. The aforementioned factor of stochastic dependence in MGF-bounds leads to a big advantage towards the PMOO-SFA, as we show in a more detailed way in Section 7.3.

Remark 2.36. The inequalities we apply as well as the order of operations is very similar for all performance bounds. This is the reason why we move the computational details for the following examples to the appendix.

Proposition 2.37 (Delay Bound with SFA)**.** *For the delay bounds, we obtain with the SFA*

$$
\begin{aligned}
\mathrm{P}(d(t)>T) &\overset{(2.13)}{\le} \mathrm{E}\Big[e^{\theta A\oslash S\,(t+T,t)}\Big]\\
&\dots \le \sum_{k=0}^{t+T}\left(E\left[e^{\theta A_1(k,t)}\right]\sum_{i=k}^{t+T}E\left[e^{p\theta A_2(k,i)}\right]^{\frac{1}{p}}E\left[e^{-p\theta S_1(k,i)}\right]^{\frac{1}{p}}\right.
\end{aligned}
$$

$$\cdot \left(\sum_{j=0}^{i} E\left[e^{q\theta A_2(j,t)}\right]^{\frac{1}{q}} E\left[e^{-q\theta S_1(j,i)}\right]^{\frac{1}{q}} \right) E\left[e^{-q\theta S_2(i,t+T)}\right]^{\frac{1}{q}} \Bigg),$$

with

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Proof. For the computational details, see A.1 in the appendix. □

Proposition 2.38 (Delay Bound with PMOO-SFA)**.** *The PMOO yields for the delay bound*

$$\begin{aligned} \mathrm{P}(d(t) > T) &\leq \sum_{k=0}^{t+T} \mathrm{E}\Big[e^{\theta(A_{\mathrm{foi}}(k,t) - S_{\mathrm{l.o.}}(k,t+T))}\Big] \\ \ldots &\leq \sum_{k=0}^{t+T} \left(E\left[e^{\theta A_1(k,t)}\right] E\left[e^{\theta A_2(k,t+T)}\right] \sum_{i=k}^{t+T} E\left[e^{-\theta S_1(k,i)}\right] E\left[e^{-\theta S_2(i,t+T)}\right] \right). \end{aligned}$$

Proof. Again, for the details check Appendix A.2. □

2.7. n-Chain Tandem

We extend the setting in Example 2.6 to the one depicted in Figure 2.5. It is the most extreme scenario one could think of in a comparison between SFA and PMOO-SFA. The fact that all the cross flows share the same servers as the foi makes it perfect for the servers to be convolved at first (strategy by PMOO) whereas the SFA causes additional dependencies at each server, as we see in this section.

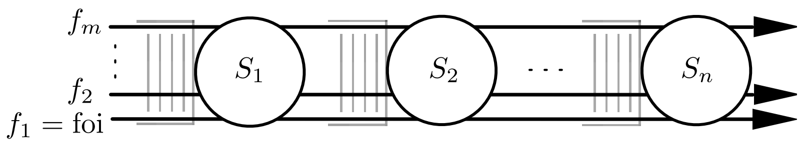


Figure 2.5.: n-chain tandem.

The SFA yields for the leftover service

$$\begin{aligned} S_{\text{leftover}} = &\left[S_1 - \sum_{j=2}^{m} A_j\right]^+ \otimes \left[S_2 - \left(\sum_{j=2}^{m} A_j\right) \oslash S_1\right]^+ \otimes \\ &\cdots \otimes \left[S_n - \left(\left(\left(\sum_{j=2}^{m} A_j\right) \oslash S_1\right) \ldots\right) \oslash S_{n-1}\right]^+, \end{aligned}$$

i.e. we subtract the cross flows, deconvolve them by the proceeding servers and convolute all the servers in the end.

Proposition 2.39 (Output Bound with SFA)**.** *For the output bound, this yields*

$$\begin{aligned}
\phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\leq} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o}}(s,t)}\right] \\
\cdots &\leq \sum_{k_0=0}^{s} \left(\mathrm{E}\left[e^{\theta A_1(k_0,t)}\right] \sum_{k_0 \leq k_1 \leq s} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq s} E\left[e^{p_1 \theta \sum_{j=2}^{m} A_j(k_0,k_1)} e^{-p_1 \theta S_1(k_0,k_1)}\right]^{\frac{1}{p_1}} \right. \\
&\left. \cdots E\left[e^{p_n \theta \left(\left(\left(\left(\sum_{j=2}^{m} A_j \right) \oslash S_1 \right) \cdots \right) \oslash S_{n-1} \right)(k_{n-1},s)} e^{-p_n \theta S_n(k_{n-1},s)}\right]^{\frac{1}{p_n}} \right),
\end{aligned}$$

with

$$\sum_{i=1}^{n} \frac{1}{p_i} = 1.$$

Proof. This time, the computation was quite involved after an application of the generalized Hölder inequality leading to a quite complicated computation as one sees in Appendix A.3. □

For the PMOO-SFA, the result is much easier. Due to the advantage of the "convolute first"-strategy for this topology, we arrive at

$$S_{\mathrm{l.o.}} = \left[\bigotimes_{i=1}^{n} S_i - \sum_{j=2}^{m} A_j \right]^{+}$$

with

$$\bigotimes_{i=1}^{n} S_i \coloneqq S_1 \otimes \cdots \otimes S_n,$$

After applying some well-known inequalities (see Section 2.5) we get the following proposition.

Proposition 2.40 (Output Bound with PMOO-SFA)**.** *The PMOO-SFA yields for the output bound*

$$\begin{aligned}
\phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\leq} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o}}(s,t)}\right] \\
\cdots &\leq \sum_{k_0=0}^{s} \left(\mathrm{E}\left[e^{\theta A_1(k_0,t)}\right] \mathrm{E}\left[e^{\theta \sum_{j=2}^{s} A_j(k_0,s)}\right] \right. \\
&\left. \sum_{k_0 \leq k_1 \leq \cdots \leq k_{n-1} \leq s} \left(\mathrm{E}\left[e^{-\theta S_1(k_0,k_1)}\right] \cdots \mathrm{E}\left[e^{-\theta S_n(k_{n-1},s)}\right] \right) \right).
\end{aligned}$$

Proof. See Appendix A.4. □

Part II.

Moment Generating Functions for Regulated Traffic

3. Why is Regulated Traffic Interesting for the SNC?

This part is based on the work of [MB00], [CCS01] as well as [VL01]. They all have in common that their arrivals are assumed to be regulated traffic, which means that the inputs in an interval $[s, t]$ are bounded deterministically. That is we assume for the arrivals

$$A(s,t) \leq \alpha(t-s), \tag{3.1}$$

where the function α is either given directly or is defined by certain properties. The exact definitions and assumptions differ which is the reason why we treat this in separate chapters. In order to answer the title-giving question of this chapter, we would like to point out the advantages of examining regulated traffic:

- Apart from inequality (3.1), no really limiting assumptions are made which makes it attractive to model traffic.
- Assuming A to be deterministically bounded is obviously a standard assumption in the deterministic network calculus (DNC). Thus, it is perfectly suitable for directly comparing bounds to the DNC and, thus, quantifying the gain of statistical multiplexing.
- This area is, compared to other fields of network calculus, less explored leaving space for new results.

For example, [RSF07] used the bounds in [VL01, VLB02] in order to evaluate the distribution of end-to-end delays in Avionics Full Duplex Switched Ethernets (AFDX), which is used, e.g., in the Airbus A380.

In this part, we show that the tail-bounds obtained in the given literature can be transformed into MGF-bounds using the Conversion Theorem 2.9. This transformation enables us to compute MGF-bounds according to this model, as it is done in Part III.

These calculations, however, are quite involved since, as we see below, the resulting integral is not easy to solve.

At last, we also compare the models introduced in Chapters 4 and 5 as far as they are comparable and explain some of their characteristics in detail.

4. Arrivals Regulated by a Leaky Bucket

4.1. Assumptions

This chapter uses the tail-bound obtained in [MB00]. We make the following assumption on the arrivals:

Definition 4.1 (Leaky Bucket Regulator)**.** The *leaky bucket regulator* guarantees that for constants ρ and σ

$$A(s,t) \leq \rho \cdot (t-s) + \sigma \qquad s \leq t, \tag{4.1}$$

where $A(s,t)$ denotes the arrivals as in Definition 2.1, except that we additionally require A to have stationary increments. By stationary increments, we mean, that for $A(t) = \sum_{i=0}^{t} a(i)$, the $a(i), 1, \ldots t$ have the same distribution.

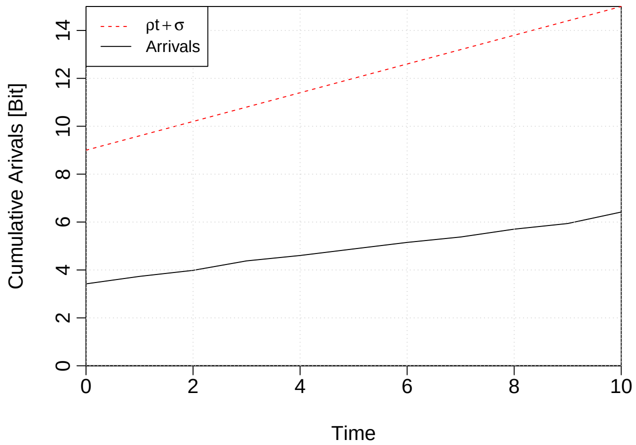


Figure 4.1.: Arrivals bounded by a leaky bucket curve.

Remark 4.2. Due to the fact that we work in this thesis on a discrete time space it already holds that

$$A(s,t) = A(s,u) + A(u,t) \qquad s \leq u \leq t, \tag{4.2}$$

which is also known as “Chasles’ relation”.

4.2. Restating the Paper’s Results

In the next step, we quote the main theorem in [MB00], Theorem 1:

Theorem 4.3. *Any stationary traffic process A which satisfies the (σ, ρ)-constraint (4.1) and the non-negativity and additivity constraints (4.2) is such that, for all convex non-decreasing functions $f : \mathbb{R} \to \mathbb{R}$, and all $t \geq 0$, one has*

$$\mathrm{E}[f\left(A(0,t) - \rho t\right)] \leq \frac{1}{2}\left(f\left(\sigma\right) + f\left(-\sigma\right)\right). \tag{4.3}$$

Proof. See [MB00]. □

Following Lemma 2 in [MB00], the authors obtained with $A := \sum_{i=1}^{n} A_i$ (in this part, A is mostly an aggregate of mutually independent traffic processes A_i), $\rho = \sum_{i=1}^{n} \rho_i$ and $\Sigma = \sum_{i=1}^{n} \sigma_i$

$$\begin{aligned}
\mathrm{P}(A(s,t) > \rho(t-s) + \gamma) &\overset{(2.3)}{\leq} \inf_{\theta \geq 0}\left\{e^{-\gamma\theta}\,\mathrm{E}\Big[e^{\theta(A(s,t)-\rho(t-s))}\Big]\right\} \\
&= \inf_{\theta \geq 0}\left\{e^{-\gamma\theta}\,\mathrm{E}\Big[e^{\theta\left(\sum_{i=1}^{n} A_i(s,t) - \sum_{i=1}^{n} \rho_i(t-s)\right)}\Big]\right\} \\
&\overset{\text{(indep.)}}{=} \inf_{\theta \geq 0}\left\{e^{-\gamma\theta}\prod_{i=1}^{n}\mathrm{E}\Big[e^{\theta(A_i(s,t)-\rho_i(t-s))}\Big]\right\} \\
&\overset{(4.3)}{\leq} \inf_{\theta \geq 0}\left\{e^{-\gamma\theta}\prod_{i=1}^{n}\frac{1}{2}\left(e^{\theta\sigma_i} + e^{-\theta\sigma_i}\right)\right\} \\
&= \inf_{\theta \geq 0}\left\{e^{-\gamma\theta}\prod_{i=1}^{n}\cosh\left(\theta\sigma_i\right)\right\},
\end{aligned} \tag{4.4}$$

where the authors of [MB00] applied Theorem 4.3 with $f\left(\sigma\right) = e^{\theta\sigma}$, which is convex and increasing.

Remark 4.4. For $n = 1$, we can compute the bound explicitly. We define

$$M(\theta) := e^{-\gamma\theta} \cdot \frac{1}{2}\left(e^{\theta\sigma} + e^{-\theta\sigma}\right) = \frac{1}{2}\exp\left\{\theta\left(\sigma - \gamma\right)\right\} + \frac{1}{2}\exp\left\{\theta\left(-\sigma - \gamma\right)\right\}. \tag{4.5}$$

The first derivative yields

$$\begin{aligned}
&\tfrac{\partial M(\theta)}{\partial\theta} = \tfrac{1}{2}\left(\sigma - \gamma\right)\exp\left\{\theta\left(\sigma-\gamma\right)\right\} + \tfrac{1}{2}\left(-\sigma-\gamma\right)\exp\left\{\theta\left(-\sigma-\gamma\right)\right\} \overset{!}{=} 0 \\
\Leftrightarrow \quad &\left(\sigma-\gamma\right)\exp\left\{\theta\left(\sigma-\gamma\right)\right\} = \left(\sigma+\gamma\right)\exp\left\{\theta\left(-\sigma-\gamma\right)\right\}
\end{aligned}$$

If $\gamma > \sigma$, then the left hand side is ≤ 0 and since the right hand side is ≥ 0 we can conclude that equality can only hold if $\theta = \infty$. So we continue with $\gamma < \sigma$. This leads to

$$\begin{aligned}
&\left(\sigma-\gamma\right)\exp\left\{\theta\left(\sigma-\gamma\right)\right\} &&= \left(\sigma+\gamma\right)\exp\left\{\theta\left(-\sigma-\gamma\right)\right\} \\
\Leftrightarrow \quad &\exp\left\{\theta\left(\sigma-\gamma\right) - \theta\left(-\sigma-\gamma\right)\right\} &&= \frac{\sigma+\gamma}{\sigma-\gamma} \\
\Leftrightarrow \quad &e^{2\theta\sigma} &&= \frac{\sigma+\gamma}{\sigma-\gamma} \\
\Leftrightarrow \quad &\theta &&= \frac{1}{2\sigma}\log\left\{\frac{\sigma+\gamma}{\sigma-\gamma}\right\}.
\end{aligned}$$

The fact that we get for the second derivative (by using that $\sigma > \gamma$)

$$\begin{aligned}\frac{\partial^2 M(\theta)}{\partial \theta^2} &= \frac{1}{2}(\sigma-\gamma)^2 \exp\{\theta(\sigma-\gamma)\} + \frac{1}{2}(-\sigma-\gamma)^2 \exp\{\theta(-\sigma-\gamma)\} \\ &> 0,\end{aligned}$$

tells us that we computed the minimum.

We insert $\theta_0 = \frac{1}{2\sigma}\log\left\{\frac{\sigma+\gamma}{\sigma-\gamma}\right\}$ into (4.5) and get

$$\begin{aligned}M(\theta_0) &\overset{(4.5)}{=} e^{-\gamma\theta_0} \cdot \frac{1}{2}\left(e^{\theta_0\sigma} + e^{-\theta_0\sigma}\right) \\ &= \left(\frac{\sigma+\gamma}{\sigma-\gamma}\right)^{-\frac{\gamma}{2\sigma}} \cdot \frac{1}{2}\left(\sqrt{\left(\frac{\sigma+\gamma}{\sigma-\gamma}\right)} + \frac{1}{\sqrt{\left(\frac{\sigma+\gamma}{\sigma-\gamma}\right)}}\right) \\ &= \frac{1}{2}\left(\left(\frac{\sigma+\gamma}{\sigma-\gamma}\right)^{-\frac{\gamma}{2\sigma}+\frac{1}{2}} + \left(\frac{\sigma+\gamma}{\sigma-\gamma}\right)^{-\frac{\gamma}{2\sigma}-\frac{1}{2}}\right).\end{aligned}$$

Instead of using the bound in (4.4), we prefer the bound of Lemma 2 in [MB00] that concluded

$$\mathrm{P}(A(s,t) > \rho(t-s)+\gamma) \leq \exp\left\{-\frac{\gamma^2}{2\sum_{i=1}^{n}\sigma_i^2}\right\}, \tag{4.6}$$

which can be interpreted as an application of the Hoeffding - inequality ([Hoe63]). In this inequality, there is no θ to be optimized anymore.

4.3. Transformation to an MGF-Bound

4.3.1. Application of the Conversion Theorem

As we already described in the motivation (beginning of chapter 3), our goal is to transform the tail-bound given in (4.6) into an MGF-bound by applying the Conversion Theorem 2.9. Owing to the fact that it needs a tail-bound $\mathrm{P}(A(s,t) > \alpha(t-s,\varepsilon)) \leq \eta(t-s,\varepsilon)$ with error $\eta(t,\varepsilon) = \varepsilon$, we set

$$\varepsilon \overset{!}{=} \eta(t,\gamma) \overset{(4.6)}{=} \exp\left\{-\frac{\gamma^2}{2\sum_{i=1}^{n}\sigma_i^2}\right\}. \tag{4.7}$$

Solving (4.7) for γ yields

$$\begin{aligned} && \varepsilon &= \exp\left\{-\frac{\gamma^2}{2\sum_{i=1}^n \sigma_i^2}\right\} \\ \Leftrightarrow && \log(\varepsilon) &= -\frac{\gamma^2}{2\sum_{i=1}^n \sigma_i^2} \\ \Leftrightarrow && -\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2 &= \gamma^2 \\ \Leftrightarrow && \gamma &= \sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}. \end{aligned}$$

We use this to reformulate α to express the dependence of ε and have

$$\alpha(t-s,\varepsilon) \overset{(4.6)}{=} \rho\cdot(t-s)+\gamma = \rho\cdot(t-s)+\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}.$$

Applying the Conversion Theorem gives us

$$\begin{aligned} \phi_{A(s,t)}(\theta) &\overset{(2.4)}{\leq} \int_0^1 e^{\theta\alpha(t-s,\varepsilon)}\mathrm{d}\varepsilon \\ &= \int_0^1 e^{\theta(\rho\cdot(t-s)+\gamma(\varepsilon))}\mathrm{d}\varepsilon \\ &= e^{\theta\rho\cdot(t-s)}\int_0^1 e^{\theta\gamma(\varepsilon)}\mathrm{d}\varepsilon \\ &= e^{\theta\rho\cdot(t-s)}\int_0^1 \underbrace{e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}}}_{=:f_1(\varepsilon)}\mathrm{d}\varepsilon. \end{aligned}$$

Solving this integral is quite involved, as we see below.

4.3.2. Solving the Integral

We insert the integral $\int_0^1 f_1(\varepsilon)\mathrm{d}\varepsilon$ in an online integral solving tool[1] and receive

$$\begin{aligned} \int_0^1 f_1(\varepsilon)\mathrm{d}\varepsilon &= \lim_{z\to 0}\left\{\varepsilon \mathrm{e}^{\sqrt{2}\sqrt{\sum_{i=1}^n \sigma_i^2}\theta\sqrt{-\log(\varepsilon)}} - \frac{\sqrt{\pi}\sqrt{\sum_{i=1}^n \sigma_i^2}\theta \mathrm{e}^{\frac{\sum_{i=1}^n \sigma_i^2\theta^2}{2}}\operatorname{erf}\left(\frac{\sqrt{2}\sqrt{-\log(\varepsilon)}-\sqrt{\sum_{i=1}^n \sigma_i^2}\theta}{\sqrt{2}}\right)}{\sqrt{2}}\right\}\Bigg|_z^1 \\ &= \lim_{z\to 0}\left\{\varepsilon e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}} - \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n \sigma_i^2}\operatorname{erf}\left(\sqrt{-\log(\varepsilon)}-\theta\sqrt{\frac{1}{2}\sum_{i=1}^n \sigma_i^2}\right)\right\}\Bigg|_z^1 \end{aligned}$$

where erf(.) denotes the *Gauss error function*

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\mathrm{d}t.$$

[1] http://www.integral-calculator.com

It holds true that

$$\lim_{x\to-\infty} \operatorname{erf}(x) = -1, \quad \lim_{x\to\infty} \operatorname{erf}(x) = 1$$

as well as

$$\operatorname{erf}(-x) = -\operatorname{erf}(x)\,.$$

At first glance, it seems like the solution came out of nowhere. But the tool also provides us with the exact steps of the calculation that we checked in Appendix B.

For the minuend, we obtain

$$\lim_{\varepsilon\searrow 0} \varepsilon e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}} = 0 \tag{4.8}$$

by using Wolfram Alpha[2].

Thus, we have

$$\begin{aligned}
\int_0^1 f_1(\varepsilon)\mathrm{d}\varepsilon &= \lim_{z\to 0}\left\{\varepsilon e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}} + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2} \operatorname{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n\sigma_i^2} - \sqrt{-\log(\varepsilon)}\right)\right\}\Bigg|_z^1 \\
&= \left[1 + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2} \operatorname{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n\sigma_i^2}\right)\right] \\
&\quad - \left[0 + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2} \cdot (-1)\right] \\
&= 1 + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2}\left(\underbrace{\operatorname{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n\sigma_i^2}\right)}_{\le 1} + 1\right) &\text{(4.9)} \\
&\le 1 + 2\sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2} = 1 + \sqrt{2\pi\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2}. &\text{(4.10)}
\end{aligned}$$

Thanks to this step, we solved and upper bounded the integral which leads us to the following MGF-bounds:

$$\begin{aligned}
\phi_{A(s,t)}(\theta) &\overset{(4.9)}{\le} e^{\theta\rho\cdot(t-s)}\left(1 + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2}\underbrace{\left(\operatorname{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n\sigma_i^2}\right) + 1\right)}_{\le 2}\right) &\text{(4.11)} \\
&\overset{(4.10)}{\le} e^{\theta\rho\cdot(t-s)}\left(1 + \sqrt{2\pi\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n\sigma_i^2}\right). &\text{(4.12)}
\end{aligned}$$

[2] http://www.wolframalpha.com/input/?i=lim(epsilon*exp(theta*sqrt(-ln(epsilon)*2*sigma)),+epsilon+to+0)

Now, we are able to compute MGF-bounds with the model introduced in [MB00]. In Chapter 6, we take a look on how these bounds perform numerically.

5. Arrivals Regulated by a Subadditive Function

5.1. Assumptions

This chapter is based on the tail-bound in [VL01]. Similar to the procedure in Chapter 4, we restate the authors' assumptions as well as their main results and transform the given bound into an MGF - version. In the next chapter, we show the results of the bounds' implementations.

(A1) The arrivals A are, as in the previous chapter, an aggregate, i.e., $A(s,t) = \sum_{i=1}^{n} A_i(s,t)$ with A_i, $i = 1, \ldots, n$ being independent.

(A2) Only the existence of an arrival curve for the arrival flow is assumed, i.e.

$$A_i(s,t) \leq \alpha_i(t-s) \quad \mathrm{Pr}-\text{a.s.}$$

which is more general than the leaky bucket assumption (4.1). We require α_i to be nonnegative, non-decreasing and $\alpha_i(t) = 0$ for $t < 0$. Additionally, we assume that α_i is subadditive, i.e., $\alpha_i(t+s) \leq \alpha_i(t) + \alpha_i(s)$ for all $s, t \in \mathbb{R}$.

(A3) For each $i = 1, \ldots, n$ and $s, t \in \mathbb{R}$,

$$\mathrm{E}[A_i(s,t)] \leq \bar{\alpha}_i \cdot (t-s)$$

where

$$\bar{\alpha}_i := \lim_{t \to \infty} \frac{\alpha_i(t)}{t}.$$

(A4) There exists a sequence of random points ("construction points")

$$\cdots < S_{-2} < S_{-1} < S_0 \leq 0 < S_1 < S_2 < \ldots$$

such that $\lim_{n \to -\infty} S_n = -\infty$ and $\lim_{n \to \infty} S_n = \infty$, and for all $n \in \mathbb{Z}$, $A(S_n, S_{n+1}) = A'(S_n, S_{n+1})$ Pr-a.s.

(A5) Define $\mathcal{S}(t) = \{S_n, n \in \mathbb{Z} : S_n \leq t\}$. The network element offers the service curve β to the aggregate of all flows, if for all $t \in \mathbb{R}$ and $u \in \mathcal{S}(t)$

$$\exists s \in [u,t]: \quad A'(u,t) - A(u,s) \geq \beta(t-s), \quad \mathrm{Pr}-\text{a.s.}$$

where β is a nonnegative non-decreasing function.

Remark 5.1. The subadditivity in (A2) exhibits some distinct similarities to our leaky bucket assumption (4.1) in Chapter 4: It means that the "evil" bursts can only occur at the beginning since the slope is decreasing.

5.2. Restating the Paper's Results

We apply the bounds obtained in [VL01]. Let

$$q(t) := A(s,t) - A'(s,t) \tag{5.1}$$

be the backlog where A' denotes the output and $s \in \mathcal{S}(t)$. We assume the service to be linear, i.e.,

$$\beta(t) = r \cdot t$$

with $r \geq \rho \geq 0$ (stability condition). Thus, we can require that

$$A'(s,t) \leq \beta(t-s) = r \cdot (t-s). \tag{5.2}$$

Then we have

$$\begin{aligned} \mathrm{P}(A(s,t) > r \cdot (t-s) + \gamma) &\overset{(5.2)}{\leq} \mathrm{P}(A(s,t) > A'(s,t) + \gamma) \\ &= \mathrm{P}(A(s,t) - A'(s,t) > \gamma) \\ &\overset{(5.1)}{=} \mathrm{P}(q(t) > \gamma)\,. \end{aligned} \tag{5.3}$$

Due to the fact that [VL01] gives us bounds on the probability that the backlog exceeds a certain value γ (called q in their paper), we have now shown by (5.3) that we can use this to obtain tail-bounds with $\alpha\,(t-s,\gamma) = r \cdot (t-s) + \gamma$.

Definition 5.2 (Vertical and Horizontal Deviation)**.** For two functions f and g, we define the *vertical deviation* by

$$v\,(f,g) = \sup_{t \geq 0} \{f(t) - g(t)\} \tag{5.4}$$

and the *horizontal deviation* by

$$h\,(f,g) = \sup_{t \geq 0} \{\inf\{u \geq 0 : f(t) \leq g(t+u)\}\}\,. \tag{5.5}$$

Remark 5.3. Figure 5.1 explains why we call it horizontal and vertical deviation. Let the arrivals $A(t)$ be upper bounded by the affine curve $\alpha(t)$ (affine because of the possible burst at the beginning) and let the service $S(t)$ be lower bounded by $\beta(t)$ (which is a rate-latency server as in Example 2.12).

Then we see that $v\,(\alpha,\beta)$ exactly resembles the vertical distance between α and β since the slope of β is bigger than α's and so the maximal distance is right at the end of β's latency. The vertical distance is also the "worst-case backlog for a network element that offers the service curve g to the aggregate arrival process that has f as an arrival curve".

On the other hand, we observe why $h\,(\alpha,\beta)$ is the horizontal distance: It is the maximal time (over all t) it took for β to fulfill $\alpha(t) = \beta\,(t + h\,(\alpha,\beta))$ (in the definition, we have $f(t) \leq g(t+u)$ since the functions are not necessarily continuous). It is also the "worst-case virtual delay" (=worst case delay, if first-in-first-out (FIFO) can be assumed for the servers).

The plot is also able to grasp one of the main problems of network calculus: Due to the fact that neither the arrival curve nor the service curve are very close to the actual traffic, the horizontal and vertical deviation "deviate" a lot by their true values. So we see that this simplification of switching to bounds might lead to a strong increase in inaccuracy.

Definition 5.4. We define

$$\bar{\alpha} := \sum_{i=1}^{n} \bar{\alpha}_i$$

and

$$\alpha := \sum_{i=1}^{n} \alpha_i.$$

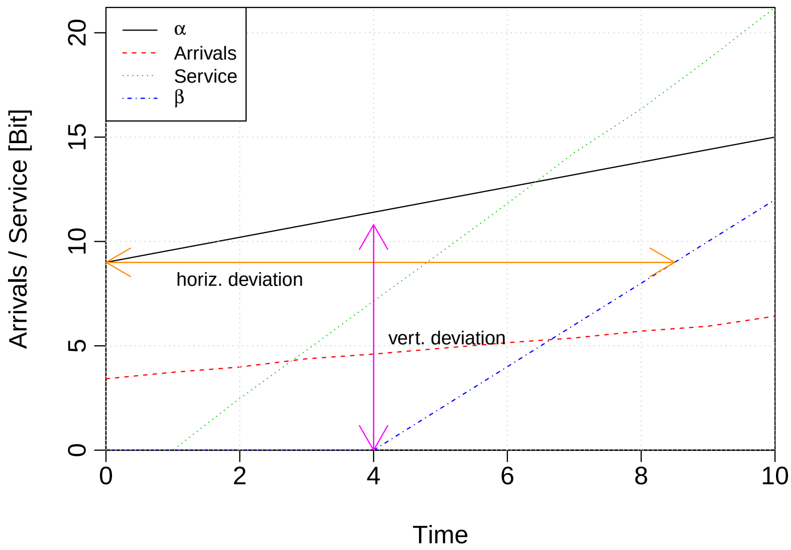


Figure 5.1.: Horizontal and vertical deviation.

In the following, we claim that in Theorem 2 in [VL01], we could have replaced $q(0)$ by $q(t)$ giving us a more general version.

Theorem 5.5. *Let $\mathcal{G} = \{g_1, \ldots, g_n\} \in \mathbb{R}_+^n : \forall i = 1, \ldots, n, \sum_{i=1}^{n} g_i \leq 1$. Assume, in addition, that for each $i = 1, \ldots, n$, (A4) holds for a virtual node that offers the service curve $g_i\beta$ fed with the arrival process A_i. Then, for any $\underline{g} \in \mathcal{G}$, and $\sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i) < \gamma < v(\alpha, \beta)$*

$$\mathrm{P}(q(t) > \gamma) \leq \exp\left\{-\frac{2\left(\gamma - \sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}\right\}. \tag{5.6}$$

Proof. We can just repeat the proof in [VL01] with $q(t)$ instead of $q(0)$ since the given inequalities for $q(t)$ hold for all t. For the details, check C in the appendix. $\square$

If we assume for the arrival that $\alpha_i(t) = \rho_i \cdot t + \sigma_i$ (leaky bucket regulator), it will hold that

$$\bar{\alpha}_i = \lim_{t\to\infty} \frac{\alpha_i(t)}{t} \overset{(4.1)}{=} \lim_{t\to\infty} \frac{\rho_i \cdot t + \sigma_i}{t} = \lim_{t\to\infty} \rho_i + \frac{\sigma_i}{t} = \rho_i.$$

Non-homogeneous case Requiring all flows to be bounded by the same arrival curve is quite limiting. Since we want to keep the amount of assumptions as small as possible, we focus on the non-homogeneous case in Theorem 2 in [VL01] instead of applying the tighter bound given in Theorem 1.

Vertical and horizontal distance For the sake of simplicity as well as generality, we just continue with h and v instead of inserting the assumptions of a leaky bucket and a constant service rate. Nonetheless, we do not want this case to be withheld:

$$\begin{aligned} v\left(\alpha_i, g_i\beta\right) &\overset{(5.4)}{=} \sup_{t>0} \left\{\alpha_i(t) - g_i\beta(t)\right\} \\ &\overset{(4.1)}{=} \sup_{t>0} \left\{\rho_i \cdot t + \sigma_i - g_i r t\right\} \\ &= \sigma_i + \sup_{t>0} \left\{t \cdot (\rho_i - g_i r)\right\} \in \{\infty, \sigma_i\} \end{aligned} \tag{5.7}$$

depending on the choice of $0 \leq g_i \leq 1$. It is clear that we have to require in the following that

$$g_i r > \rho_i \tag{5.8}$$

so that the vertical deviation is finite and thus

$$v\left(\alpha_i, g_i\beta\right) = \sigma_i.$$

As a consequence, we can assume $v(\alpha, \beta) = \sum_{i=1}^n \sigma_i$ in the leaky bucket / constant service rate case.

For the horizontal deviation, this means

$$\begin{aligned} h\left(\alpha_i, g_i\beta\right) &\overset{(5.5)}{=} \sup_{t\geq 0} \left\{\inf \left\{u \geq 0 : \alpha_i(t) \leq g_i\beta(t+u)\right\}\right\} \\ &= \sup_{t\geq 0} \left\{\inf \left\{u \geq 0 : \rho_i \cdot t + \sigma_i \leq g_i r \cdot (t+u)\right\}\right\} \\ &\overset{(5.8)}{=} \sup_{t\geq 0} \left\{\inf \left\{u \geq 0 : \left(\underbrace{\rho_i - g_i r}_{<0}\right) t + \sigma_i \leq g_i r u\right\}\right\} \\ &\overset{(\text{sup att. for } t=0)}{=} \sup_{t=0} \left\{\inf \left\{u \geq 0 : \sigma_i \leq g_i r u\right\}\right\} \\ &= \inf\left\{u \geq 0 : \sigma_i \leq g_i r u\right\} = \frac{\sigma_i}{g_i r}. \end{aligned} \tag{5.9}$$

5.3. Transformation to an MGF-Bound

We repeat the procedure from Section 4.3 for the results of [VL01].

5.3.1. Application of the Conversion Theorem

By inequality (5.6), we know that A is tail-bounded for $\sum_{i=1}^n \bar{\alpha}_i h(\alpha_i) < \gamma < v(\alpha, \beta)$ with envelope $\alpha(t-s, \gamma) = \rho \cdot (t-s) + \gamma$ and

$$\varepsilon \stackrel{!}{=} \eta(t, \gamma) = \exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^n v(\alpha_i, g_i\beta)^2} \right\}, \tag{5.10}$$

since

$$\mathrm{P}(A(s,t) > \rho(t-s) + \gamma) \stackrel{(5.3)}{\leq} \mathrm{P}(q(t) > \gamma) \stackrel{(5.6)}{\leq} \exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^n v(\alpha_i, g_i\beta)^2} \right\}.$$

If we assumed again the arrivals to be leaky bucket and the service to be of constant rate, i.e., $\bar{\alpha}_i = \rho_i$, $v(\alpha_i, g_i\beta) = \sigma_i$ and $h(\alpha_i, g_i\beta) = \frac{\sigma_i}{g_i r}$, this would be equivalent to

$$\mathrm{P}(A(s,t) > \rho(t-s) + \gamma) \leq \exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^n \rho_i \frac{\sigma_i}{g_i r}\right)^2}{\sum_{i=1}^n \sigma_i^2} \right\} \tag{5.11}$$

for $\sum_{i=1}^n \rho_i \frac{\sigma_i}{g_i r} < \gamma < \sum_{i=1}^n \sigma_i$. At this point, we would like to mention that this bound only holds if γ is in between these bounds, as we see below.

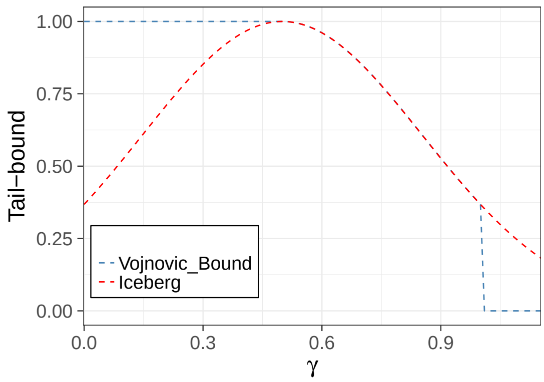
Solving (5.10) for γ yields

$$\begin{aligned}
& \varepsilon && = \exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^n v(\alpha_i, g_i\beta)^2} \right\} \\
\Leftrightarrow \quad & \log(\varepsilon) && = -\frac{2\left(\gamma - \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^n v(\alpha_i, g_i\beta)^2} \\
\Leftrightarrow \quad & -\frac{1}{2}\log(\varepsilon) \sum_{i=1}^n v(\alpha_i, g_i\beta)^2 && = \left(\gamma - \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2 \\
\Leftrightarrow \quad & \sqrt{-\frac{1}{2}\log(\varepsilon) \sum_{i=1}^n v(\alpha_i, g_i\beta)^2} && = \gamma - \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta) \\
\Leftrightarrow \quad & \gamma && = \sqrt{-\frac{1}{2}\log(\varepsilon) \sum_{i=1}^n v(\alpha_i, g_i\beta)^2} + \sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)
\end{aligned} \tag{5.12}$$

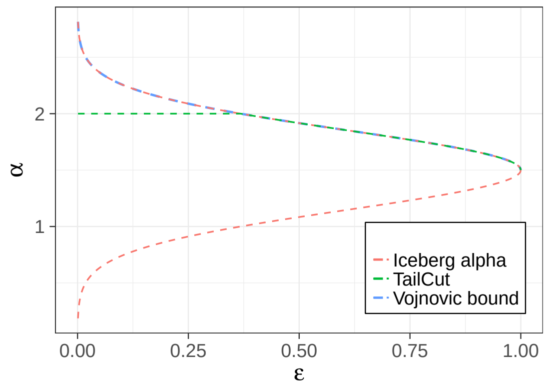
for $\sum_{i=1}^n \bar{\alpha}_i h(\alpha_i) < \gamma < v(\alpha, \beta)$. Now we see why it is necessary that γ is bounded to this interval. If we use

$$\exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^n \rho_i \frac{\sigma_i}{g_i r}\right)^2}{\sum_{i=1}^n \sigma_i^2} \right\}$$

also for $\gamma \leq \sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i)$ and $\gamma \geq v(\alpha, \beta)$, our quadratic tail becomes an “iceberg”, as we can see in Figure 5.2a.



(a) “Iceberg”-bound and normal tail-bound.



(b) “Iceberg”-α, normal α and cut α.

Figure 5.2.: “Iceberg” and normal bound in the Conversion Theorem.

This would mean that we could not invert this function, since it would not be injective anymore.

The lower bound of γ, $\sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i)$, is not violated in our integral since ε is upper bounded by 1, so we always have that $\gamma \geq \sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i)$ where equality only holds in a single point, i.e., it does not increase the integral.

We would like to mention that $\gamma < v(\alpha, \beta)$ is not fulfilled for $\varepsilon \to 0$, i.e., we overestimate the bound if we allow ε to go to 0 (as we can see in Figure 5.2b). We ignore this fact at first and come back to it later in subsection 5.3.3 (where we see that for

$$\varepsilon \leq l := \exp\left\{ \frac{-2\left(v(\alpha, \beta) - \sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2} \right\},$$

α is just

$$\alpha(t-s, \varepsilon) = \rho \cdot (t-s) + v(\alpha, \beta),$$

which is independent of ε) to show that the bound can be improved because of this issue.

We use the above to reformulate α so that we can express the dependence of ε. This gives us

$$\alpha(t-s, \varepsilon) \overset{(5.12)}{=} \rho \cdot (t-s) + \sqrt{-\frac{1}{2}\log(\varepsilon) \sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2} + \sum_{i=1}^{n} \bar{\alpha}_i h(\alpha_i, g_i\beta). \tag{5.13}$$

The leaky bucket / constant service rate assumption then leads to

$$\alpha(t-s, \varepsilon) = \rho \cdot (t-s) + \sqrt{-\frac{1}{2}\log(\varepsilon) \sum_{i=1}^{n} \sigma_i^2} + \sum_{i=1}^{n} \rho_i \frac{\sigma_i}{g_i r}.$$

The plot of α for $\varepsilon \in [0,1]$ is given on the right hand side of Figure 5.2. The lower blue line resembles the negative solution of the quadratic equation we solved before, hence we obtain a "real reflection on the first bisecting" of the left plot.

We apply the Conversion Theorem and get

$$\begin{aligned}
\phi_{A(s,t)}(\theta) &\overset{(2.4)}{\leq} \int_0^1 e^{\theta\alpha(t-s,\varepsilon)}\mathrm{d}\varepsilon \\
&= \int_0^1 e^{\theta(\rho\cdot(t-s)+\gamma(\varepsilon))}\mathrm{d}\varepsilon \\
&= e^{\theta\rho\cdot(t-s)} \int_0^1 e^{\theta\gamma(\varepsilon)}\mathrm{d}\varepsilon \\
&\overset{(5.13)}{=} e^{\theta\rho\cdot(t-s)} \int_0^1 \exp\left\{\theta\left(\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2} + \sum_{i=1}^n \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)\right)\right\}\mathrm{d}\varepsilon \\
&= e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^n \bar{\alpha}_i h(\alpha_i,g_i\beta)} \int_0^1 \underbrace{e^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}}}_{=:f_2(\varepsilon)}\mathrm{d}\varepsilon.
\end{aligned}$$

As a quick reminder, we would like to point out again that we overestimate the bound by allowing ε to converge to 0.

5.3.2. Solving the Integral

Again, we insert this integral into the aforementioned online integral solving tool[1] and receive

$$\begin{aligned}
\int_0^1 f_2(\varepsilon)\mathrm{d}\varepsilon = \lim_{z\to 0}&\left\{\varepsilon\mathrm{e}^{\sqrt{\frac{1}{2}\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}\theta\sqrt{-\log(\varepsilon)}}\right.\\
&\left.\left. - \frac{\sqrt{\pi}\sqrt{\frac{1}{2}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\theta\mathrm{e}^{\frac{\sum_{i=1}^n v(\alpha_i,g_i\beta)^2\theta^2}{8}}\operatorname{erf}\left(\frac{2\sqrt{-\log(\varepsilon)}-\sqrt{\frac{1}{2}\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}\theta}{2}\right)}{2}\right\}\right|_z^1 \\
= \lim_{z\to 0}&\left\{\varepsilon\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}}\right.\\
&\left.\left. -\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}{8}}\operatorname{erf}\left(\sqrt{-\log(\varepsilon)} - \theta\sqrt{\frac{1}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\right)\right\}\right|_z^1 \\
= \lim_{z\to 0}&\left\{\varepsilon\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}}\right.
\end{aligned}$$

[1] http://www.integral-calculator.com

$$+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}} \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} - \sqrt{-\log(\varepsilon)}\right)\Bigg\}\Bigg|_z^1$$

and hence

$$\begin{aligned}
\int_0^1 f_2(\varepsilon)\mathrm{d}\varepsilon = & \left[1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}} \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right)\right] \\
& - \left[0+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}} \cdot (-1)\right] \\
= & 1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}} \left(\underbrace{\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right)}_{\leq 1}+1\right) \qquad (5.14) \\
\leq & 1+\theta\sqrt{\frac{\pi}{2}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}}. \qquad (5.15)
\end{aligned}$$

This enables us to obtain for the MGF

$$\begin{aligned}
\phi_{A(s,t)}(\theta) \overset{(5.14)}{\leq} & e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\bar{\alpha}_i h(\alpha_i, g_i\beta)} \\
& \cdot\left(1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}} \underbrace{\left(\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right)+1\right)}_{\leq 2}\right) \\
\overset{(5.15)}{\leq} & e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\bar{\alpha}_i h(\alpha_i, g_i\beta)} \left(1+\theta\sqrt{\frac{\pi}{2}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}{8}}\right).
\end{aligned}$$

Assuming again the arrivals to be leaky bucket and the service to be constant (we use $\rho = \sum_{i=1}^{n}\rho_i$ to avoid any confusion):

$$\begin{aligned}
\phi_{A(s,t)}(\theta) \leq & e^{\theta(t-s)\sum_{i=1}^{n}\rho_i+\theta\sum_{i=1}^{n}\rho_i\frac{\sigma_i}{g_i r}} \left(1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2} \mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}} \left(\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right)+1\right)\right) \\
= & e^{\theta\sum_{i=1}^{n}\rho_i\cdot\left(t-s+\frac{\sigma_i}{g_i r}\right)} \left(1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2} \mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}} \underbrace{\left(\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right)+1\right)}_{\leq 2}\right)
\end{aligned}$$

$$\leq e^{\theta \sum_{i=1}^{n} \rho_i \cdot \left(t-s+\frac{\sigma_i}{g_i r}\right)} \left(1+\theta \sqrt{\frac{\pi}{2}\sum_{i=1}^{n}\sigma_i^2 \mathrm{e}^{\frac{\theta^2 \sum_{i=1}^{n}\sigma_i^2}{8}}}\right).$$

5.3.3. Cut the Tail

As mentioned in Subsection 5.3.1, we overestimated the MGF bound since we allow

$$\gamma(\varepsilon) \overset{(5.12)}{=} \sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} + \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)$$

to be greater than or equal to $v\left(\alpha, \beta\right)$. In this instance, we solve for ε and it holds

$$\begin{aligned}
& & v\left(\alpha,\beta\right) &\leq \gamma(\varepsilon) \\
&\Leftrightarrow & v\left(\alpha,\beta\right) &\leq \sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} + \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right) \\
&\Leftrightarrow & \sqrt{-\tfrac{1}{2}\log(\varepsilon)\textstyle\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} &\geq \underbrace{v\left(\alpha,\beta\right) - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)}_{\geq 0} \\
&\Leftrightarrow & -\tfrac{1}{2}\log(\varepsilon)\textstyle\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2 &\geq \left(v\left(\alpha,\beta\right) - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)\right)^2 \\
&\Leftrightarrow & \log(\varepsilon) &\leq \frac{-2\left(v\left(\alpha,\beta\right) - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)\right)^2}{\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \\
&\Leftrightarrow & \varepsilon &\leq \exp\left\{\frac{-2\left(v\left(\alpha,\beta\right) - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)\right)^2}{\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right\} =: l,
\end{aligned} \tag{5.16}$$

i.e., the bound $\gamma < v\left(\alpha, \beta\right)$ is violated if and only if $\varepsilon \leq l$.

We note that $v\left(\alpha,\beta\right) - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)$ is exactly the upper minus the lower bound of γ.

As a result, we can improve the MGF bound by making a case distinction for ε being greater or smaller than l in (5.16). We would like to mention that $l \in (0, 1)$ since $v\left(\alpha,\beta\right) > \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)$.

For $\varepsilon \leq l$ it holds that

$$\alpha(t-s,\varepsilon) = \rho\cdot(t-s) + v\left(\alpha,\beta\right),$$

since the burst cannot be greater than the maximal vertical deviation, whereas for $\varepsilon > l$ it remains that

$$\alpha(t-s,\varepsilon) = \rho\cdot(t-s) + \sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} + \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right).$$

Applying the conversion theorem yields

$$
\begin{aligned}
\phi_{A(s,t)}(\theta) &\leq \int_0^1 e^{\theta\alpha(t-s,\varepsilon)}\mathrm{d}\varepsilon \\
&= \int_0^l e^{\theta\alpha(t-s,\varepsilon)}\mathrm{d}\varepsilon + \int_l^1 e^{\theta\alpha(t-s,\varepsilon)}\mathrm{d}\varepsilon \\
&= \dots \\
&= l \cdot e^{\theta\rho\cdot(t-s)+\theta v(\alpha,\beta)} + e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^n \bar{\alpha}_i h(\alpha_i, g_i\beta)} \int_l^1 \underbrace{e^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}}}_{=:f_3(\varepsilon)}\mathrm{d}\varepsilon
\end{aligned}
$$

with

$$
\begin{aligned}
\int_l^1 f_3(\varepsilon)\mathrm{d}\varepsilon = &\Bigg\{\varepsilon \mathrm{e}^{\sqrt{\frac{1}{2}\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}\theta\sqrt{-\log(\varepsilon)}} \\
&- \frac{\sqrt{\pi}\sqrt{\frac{1}{2}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\theta \mathrm{e}^{\frac{\sum_{i=1}^n v(\alpha_i,g_i\beta)^2\theta^2}{8}} \operatorname{erf}\left(\frac{2\sqrt{-\log(\varepsilon)} - \sqrt{\frac{1}{2}\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}\theta}{2}\right)}{2}\Bigg\}\Bigg|_l^1 \\
= &\Bigg\{\varepsilon \mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}} \\
&-\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}{8}} \operatorname{erf}\left(\sqrt{-\log(\varepsilon)} - \theta\sqrt{\frac{1}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\right)\Bigg\}\Bigg|_l^1 \\
= &\Bigg\{\varepsilon \mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}} \\
&+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}{8}} \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2} - \sqrt{-\log(\varepsilon)}\right)\Bigg\}\Bigg|_l^1 \\
= &\left[1 + \theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}{8}} \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\right)\right] \\
&- \Bigg[l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}} + \theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}{8}} \\
&\cdot \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2} - \sqrt{-\log\left(l\right)}\right)\Bigg] \\
= &1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}} + \theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n v(\alpha_i,g_i\beta)^2}{8}}
\end{aligned}
$$

$$\cdot \left[\underbrace{\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right) - \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} - \sqrt{-\log\left(l\right)}\right)}_{\leq 2}\right] \tag{5.17}$$

$$\leq 1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}} + \theta\sqrt{\frac{\pi}{2}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}{8}} \tag{5.18}$$

where $l \overset{(5.16)}{=} \exp\left\{\frac{-2\left(v(\alpha,\beta)-\sum_{i=1}^{n}\bar{\alpha}_i h(\alpha_i, g_i\beta)\right)^2}{\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}\right\}$. Subsequently, it follows for the MGF-bound:

$$\begin{aligned}
\phi_{A(s,t)}(\theta) &\overset{(5.17)}{\leq} l \cdot e^{\theta\rho\cdot(t-s)+\theta v(\alpha,\beta)} + e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\bar{\alpha}_i h(\alpha_i, g_i\beta)} \\
&\quad \cdot \left(1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}} + \theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}{8}}\right. \\
&\quad \left.\cdot \left[\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right) - \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} - \sqrt{-\log\left(l\right)}\right)\right]\right) \\
&\overset{(5.18)}{\leq} l \cdot e^{\theta\rho\cdot(t-s)+\theta v(\alpha,\beta)} + e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\bar{\alpha}_i h(\alpha_i, g_i\beta)} \\
&\quad \cdot \left(1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}} + \theta\sqrt{\frac{\pi}{2}\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n} v(\alpha_i, g_i\beta)^2}{8}}\right).
\end{aligned}$$

Assuming again the arrivals to be regulated by the leaky bucket and the service to be constant, with $\rho = \sum_{i=1}^{n}\rho_i$, we get

$$\begin{aligned}
\phi_{A(s,t)}(\theta) &\overset{(5.17)}{\leq} l \cdot e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\sigma_i} + e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\rho_i\frac{\sigma_i}{g_i r}}\left(1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n}\sigma_i^2}}\right. \\
&\quad \left.+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}}\left[\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right) - \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2} - \sqrt{-\log\left(l\right)}\right)\right]\right) \\
&= l \cdot e^{\theta\sum_{i=1}^{n}(\rho_i(t-s)+\sigma_i)} + e^{\theta\sum_{i=1}^{n}\rho_i\cdot\left(t-s+\frac{\sigma_i}{g_i r}\right)}\left(1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n}\sigma_i^2}}\right. \\
&\quad \left.+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}}\left[\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right) - \operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2} - \sqrt{-\log\left(l\right)}\right)\right]\right) \\
&\overset{(5.18)}{\leq} l \cdot e^{\theta\sum_{i=1}^{n}(\rho_i(t-s)+\sigma_i)} + e^{\theta\sum_{i=1}^{n}\rho_i\cdot\left(t-s+\frac{\sigma_i}{g_i r}\right)}\left(1 - l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n}\sigma_i^2}}\right.
\end{aligned}$$

$$+\theta\sqrt{\frac{\pi}{2}\sum_{i=1}^{n}\sigma_i^2 \mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}}}\right),$$

with

$$l \overset{(5.16)}{=} \exp\left\{\frac{-2\left(\sum_{i=1}^{n}\sigma_i - \sum_{i=1}^{n}\rho_i\frac{\sigma_i}{g_i r}\right)^2}{\sum_{i=1}^{n}\sigma_i^2}\right\} = \exp\left\{\frac{-2\left(\sum_{i=1}^{n}\sigma_i\left(1-\frac{\rho_i}{g_i r}\right)\right)^2}{\sum_{i=1}^{n}\sigma_i^2}\right\}.$$

In the next chapter, we plot the bounds and see if cutting the tail leads to a significant impact.

6. Comparison of the Traffic Models

In this chapter, we compare the tail-bounds obtained in [MB00] and [VL01] plus the MGF-bounds we calculated in the Sections 4.3 and 5.3.

6.1. Tail-Bounds

Let us quickly restate the tail-bound in (4.6) from [MB00]:

$$\mathrm{P}(A(s,t) > \rho(t-s) + \gamma) \leq \exp\left\{ -\frac{\gamma^2}{2\sum_{i=1}^{n} \sigma_i^2} \right\}$$

with $\rho = \sum_{i=1}^{n} \rho_i$. For $\sigma_i = \rho_i = \frac{1}{n}$ and $n \in \{1, 2, 10, 100\}$ we draw a plot for different γ on the x-axis in Figure 6.1a.

We compare this to the tail-bound obtained by Vojnović and Le Boudec (see Inequality (5.6)):

$$\mathrm{P}(q(t) > \gamma) \leq \exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)\right)^2}{\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2} \right\}$$

with $\sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i\right) < \gamma < v\left(\alpha, \beta\right)$. According to (5.11), assuming the arrivals to be leaky bucket and the service to be constant, this means

$$\mathrm{P}(A(s,t) > \rho(t-s) + \gamma) \leq \exp\left\{ -\frac{2\left(\gamma - \sum_{i=1}^{n} \rho_i \frac{\sigma_i}{g_i r}\right)^2}{\sum_{i=1}^{n} \sigma_i^2} \right\}$$

for $\sum_{i=1}^{n} \rho_i \frac{\sigma_i}{g_i r} < \gamma < \sum_{i=1}^{n} \sigma_i$.

For $\sigma_i = \rho_i = \frac{1}{n}$ as well as $g_i = \frac{1}{n}$ and $r = \max\{\rho_i\} + 3$, the results are plotted in Figure 6.1b.

The trivial deterministic bound

$$\mathrm{P}(A(s,t) > \rho(t-s) + \gamma) \leq 1_{\left\{x \,:\, x \leq \sum_{i=1}^{n} \sigma_i\right\}}(\gamma) \tag{6.1}$$

serves as our benchmark, where

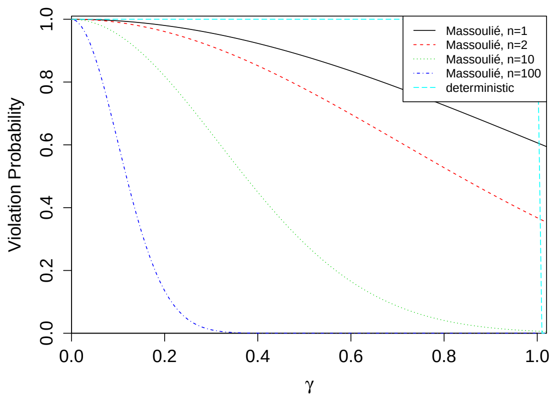
$$1_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{else} \end{cases}$$

denotes the indicator function. We made the following choice for the parameters: $\sum_{i=1}^{n} \sigma_i = 1$ and $\sum_{i=1}^{n} \rho_i = 1$. The plot in Figure 6.1c incorporates both tail-bounds, Massoulié and Vojnović, of the arrivals for different γ.

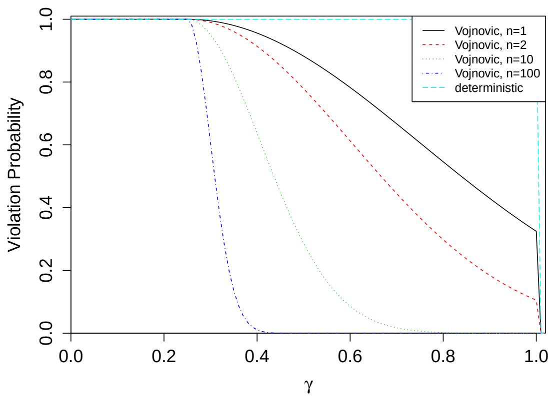
The multiplexing effect is quite evident as we fix $\sum_{i=1}^{n} \rho_i$ and $\sum_{i=1}^{n} \sigma_i$ and only increase n leading to an enforced improvement compared to the deterministic bound.

The graphs provide only limited comparability, since the bound in [VL01] includes all the parameters from [MB00] plus the g_i and the service rate r. Nonetheless, a certain "trend" is observable, namely the fact that (5.11) is worse for small γ (it is $= 1$ for $\gamma \leq \sum_{i=1}^{n} \rho_i \frac{\sigma_i}{g_i r}$), whereas the factor 2 leads to a steeper curve. This leads to better bounds for bigger γ depending on the other parameters. On the other hand, the multiplexing effect is stronger for the Massoulié bounds. We would like to allude to the fact that this is slightly influenced by the choice of g_i.

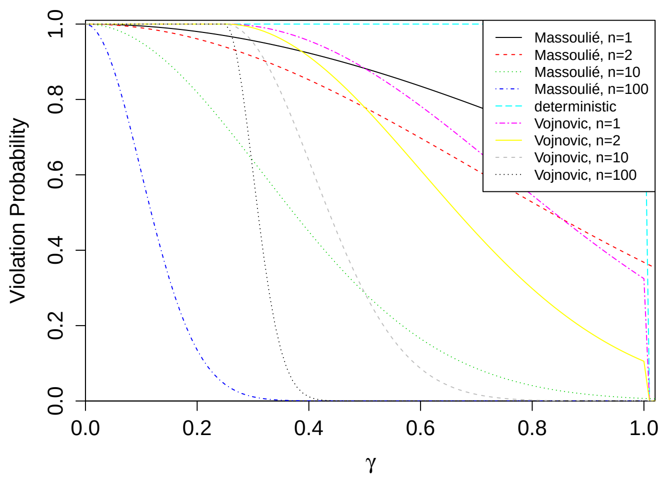
In our example, the two models can have an intersection point, e.g., for $n = 2$, since we chose the rate to be $r = \max\{\rho_i\} + 3$; if we chose r to be much bigger or smaller, it would not occur.



(a) Massoulié tail-bound.



(b) Vojnović tail-bound.



(c) Massoulié and Vojnović tail-bound merged.

Figure 6.1.: Tail-Bound Comparison.

6.2. MGF-Bounds

Now, let us compare the MGF-bounds we obtained in the Sections 4.3 and 5.3. As a reminder, we repeat them quickly: We transformed the tail-bound in [MB00] to an MGF-bound (via the Conversion Theorem 2.9) leading to

$$\begin{aligned}\phi_{A(s,t)}(\theta) &\overset{(2.4)}{\leq} e^{\theta\rho(t-s)} \int_0^1 e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}} \mathrm{d}\varepsilon \\ &\overset{(4.9)}{=} e^{\theta\rho(t-s)} \left(1+\sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n \sigma_i^2}\left(\mathrm{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n \sigma_i^2}\right)+1\right)\right).\end{aligned}$$

Similarly, we converted the tail-bound in [VL01] to

$$\begin{aligned}\phi_{A(s,t)}(\theta) &\overset{(2.4)}{\leq} e^{\theta\rho(t-s)+\theta\sum_{i=1}^n \rho_i \frac{\sigma_i}{g_i r}} \int_0^1 e^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^n v(\alpha_i, g_i\beta)^2}} \mathrm{d}\varepsilon \\ &\overset{(5.14)}{=} e^{\theta\sum_{i=1}^n \rho_i\left(t-s+\frac{\sigma_i}{g_i r}\right)} \left(1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^n \sigma_i^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^n \sigma_i^2}{8}}\left(\mathrm{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^n \sigma_i^2}\right)+1\right)\right).\end{aligned}$$

Please keep in mind that this is the overestimated version since we could have cut the tail in the integral.

Again, the trivial bound obtained by the leaky bucket assumption is used as a benchmark:

$$\phi_{A(s,t)}(\theta) = \mathrm{E}\left[e^{\theta A(s,t)}\right] \overset{(4.1)}{\leq} \mathrm{E}\left[e^{\theta(\rho(t-s)+\sigma)}\right] = \mathrm{E}\left[e^{\theta\left(\sum_{i=1}^n \rho_i(t-s)+\sum_{i=1}^n \sigma_i\right)}\right]. \tag{6.2}$$

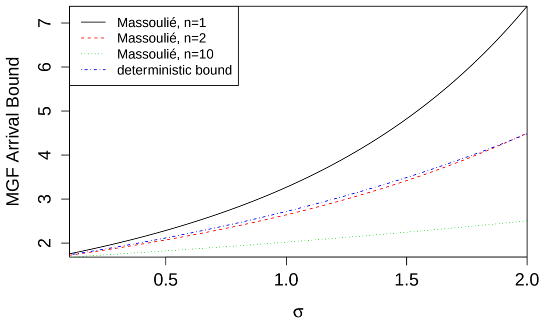
We implemented these bounds and plotted the arrivals' MGF-bounds for different σ in Figure 6.2. We select the parameters to be $\sum_{i=1}^n \rho_i = 1, t-s = 1$, and $\theta = 0.5$. Similar to Figure 6.1, the Massoulié bounds are shown in in Figure 6.2a, Vojnović bounds in Figure 6.2b and the merged plot in Figure 6.2c.

The results are quite similar to the tail-bounds, namely that the multiplexing effect is again more significant for the Massoulié bounds, i.e., a change in the number of flows, n, decreases the bound more severely. We saw that this effect could be slightly reduced by increasing the g_i but was still clearly visible.

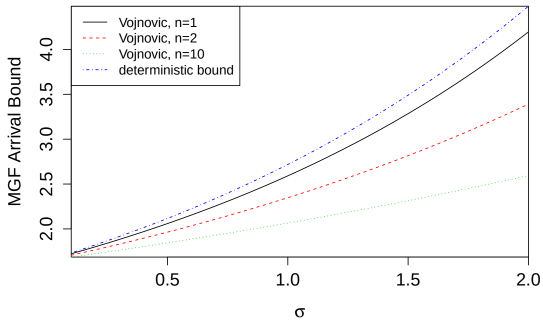
Again, we would like to mention that due to the greater parameter set in the Vojnović model, a comparability is limited.

6.3. Tail-Cuts

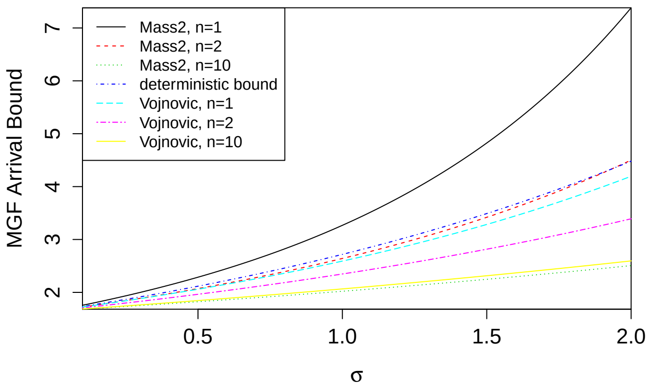
In Subsection 5.3.3, we showed that we overestimated the MGF by allowing γ to be greater than or equal to $v(\alpha, \beta)$. In this section, we would like to quantify the improvement of cutting the tail, i.e., we compare the normal MGF-bound derived from the



(a) Massoulié MGF-bound.



(b) Vojnović MGF-bound.



(c) Massoulié and Vojnović MGF-bound merged.

Figure 6.2.: MGF-bound comparison.

results of [VL01],

$$\begin{aligned}
\phi_{A(s,t)}(\theta) &\overset{(2.4)}{\leq} e^{\theta\rho(t-s)+\theta\sum_{i=1}^{n}\rho_i\frac{\sigma_i}{g_i r}} \int_0^1 e^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^{n}v(\alpha_i,g_i\beta)^2}}\mathrm{d}\varepsilon \\
&\overset{(5.14)}{=} e^{\theta\sum_{i=1}^{n}\rho_i\left(t-s+\frac{\sigma_i}{g_i r}\right)}\left(1+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}}\left(\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right)+1\right)\right),
\end{aligned}$$

and the improved bound by cutting the tail,

$$\begin{aligned}
\phi_{A(s,t)}(\theta) &\overset{(2.4)}{\leq} l\cdot e^{\theta\rho\cdot(t-s)+\theta v(\alpha,\beta)} + e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\bar{\alpha}_i h(\alpha_i,g_i\beta)}\int_l^1 e^{\theta\sqrt{-\frac{1}{2}\log(\varepsilon)\sum_{i=1}^{n}v(\alpha_i,g_i\beta)^2}}\mathrm{d}\varepsilon \\
&\overset{(5.17)}{\leq} l\cdot e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\sigma_i} + e^{\theta\rho\cdot(t-s)+\theta\sum_{i=1}^{n}\rho_i\frac{\sigma_i}{g_i r}}\Bigg(1-l\mathrm{e}^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n}\sigma_i^2}}
\end{aligned}$$

$$+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}}\left[\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right)-\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}-\sqrt{-\log\left(l\right)}\right)\right]\Bigg)$$

$$=l\cdot e^{\theta\sum_{i=1}^{n}(\rho_i(t-s)+\sigma_i)}+e^{\theta\sum_{i=1}^{n}\rho_i\cdot\left(t-s+\frac{\sigma_i}{g_i r}\right)}\Bigg(1-le^{\theta\sqrt{-\frac{1}{2}\log(l)\sum_{i=1}^{n}\sigma_i^2}}$$

$$+\theta\sqrt{\frac{\pi}{8}\sum_{i=1}^{n}\sigma_i^2}\mathrm{e}^{\frac{\theta^2\sum_{i=1}^{n}\sigma_i^2}{8}}\left[\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}\right)-\operatorname{erf}\left(\theta\sqrt{\frac{1}{8}\sum_{i=1}^{n}\sigma_i^2}-\sqrt{-\log\left(l\right)}\right)\right]\Bigg)$$

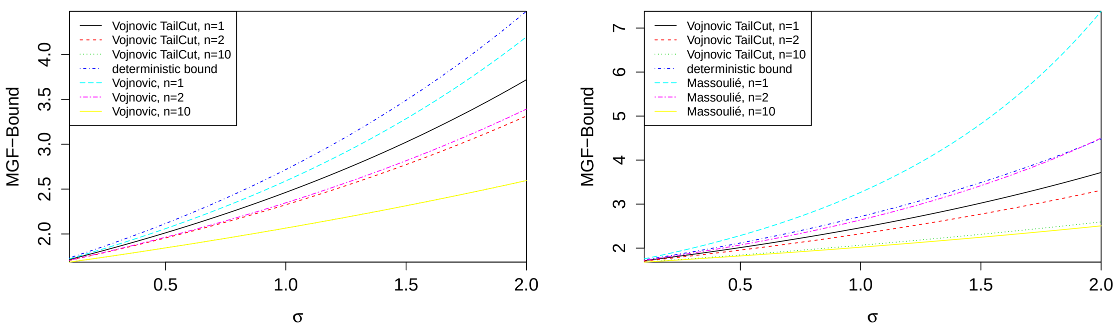
with

$$l\overset{(5.16)}{=}\exp\left\{\frac{-2\left(\sum_{i=1}^{n}\sigma_i-\sum_{i=1}^{n}\rho_i\frac{\sigma_i}{g_i r}\right)^2}{\sum_{i=1}^{n}\sigma_i^2}\right\}=\exp\left\{\frac{-2\left(\sum_{i=1}^{n}\sigma_i\left(1-\frac{\rho_i}{g_i r}\right)\right)^2}{\sum_{i=1}^{n}\sigma_i^2}\right\}.$$

With the same parameters as in Figure 6.2, we created plots for the MGF-bounds in Figure 6.3, but now for the cut tail.

We conclude by the plot in Figure 6.3a, that the improvement is rather limited, especially if we increase n and therefore the multiplexing. For example if we set σ to 2, for $n=1$, the improvement is 11.3%, but for $n=10$, it is only $7.7\cdot 10^{-5}$%.

Hence, the bound from Chapter 4 did not change significantly, as we also see in Figure 6.3b which is quite comparable to the plot in Figure 6.2c.



(a) Vojnović normal MGF-bound and with cut tail. (b) Massoulié MGF-bound and Vojnović MGF-bound with cut tail.

Figure 6.3.: Tail cut comparison.

Part III.

Analyzing Sink Trees

7. Sink Tree Performance Bounds and the Impact of Dependencies

This chapter focuses on performance bounds in sink trees as shown in Figure 7.1.

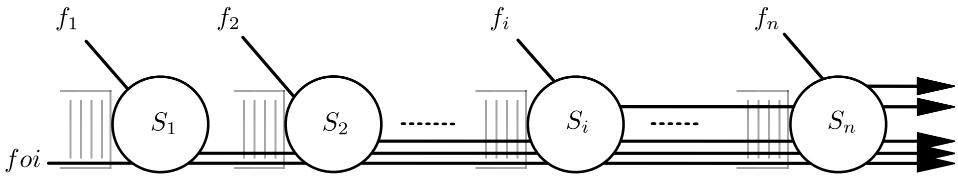


Figure 7.1.: Sink tree.

The reader might have a different expectation of what a sink tree looks like. The reason is that we only focus on the flow of interest (foi) and its perspective. From the foi's point of view, there are just the crossing arrivals and each arrival's priority.

Sink trees have various applications as in, e.g., Multi Protocol Label Switching (MPLS), where sink tree aggregation is necessary [RVC01, LMMS06], or in wireless sensor networks [SR05, KAT06].

On the other hand, more general topologies are not very well understood so far in stochastic network calculus, as we mentioned in Chapter 1. For sink trees, we are able to describe the networks in a general form due to the regularity in their shape. Additionally, as we see in this chapter, important factors such as dependencies are already quite involved here, hence we leave more general topologies for future work.

We assume that the priority is according to the representation in Figure 7.1: The flows on top have the highest priority and the priority is decreasing with their order in the figure.

The outlook of this chapter is as follows:

1. First, we make assumptions on the in-/dependence of the cross flows.

2. Then, we calculate the delay bound either with the SFA or the PMOO-SFA.

3. At last, we try to provide a numerical comparison shown in different plots. Therefore, we assume henceforth the **service** to be of **constant rate** and the MGF to

be of the form

$$\phi_{A(s,t)}(\theta) \overset{(2.4)}{\leq} e^{\theta\rho(t-s)} \int_0^1 e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}} \mathrm{d}\varepsilon$$
$$\overset{(4.9)}{=} e^{\theta\rho(t-s)} \left(1 + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2} \left(\operatorname{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n \sigma_i^2}\right) + 1\right)\right),$$

i.e., as in the paper by **Massoulié converted to MGF-bounds** as we have shown in Section 4.3.

4. Each plot also provides a version with a standard choice for the parameter $\theta = 1$ and $p_i = m$, where m is the number of dependent factors in the generalized Hölder inequality, as well as an "optimized" version. The latter one contains a systematic "grid search", where the function values are computed for θ, p_i in a 0.5-grid between 0.5, 1.5 and 4 where the lowest value is chosen.

Remark 7.1. We would like to point out that the choice of the Hölder parameter, $p_i = m$ in step 4 is always feasible, since for m dependent factors in the generalized Hölder inequality (2.22)

$$\mathrm{E}\left[\prod_{i=1}^m |X_i|\right] \overset{(2.22)}{\leq} \prod_{i=1}^m \mathrm{E}[|X_i|^{p_i}]^{\frac{1}{p_i}} = \prod_{i=1}^m \mathrm{E}[|X_i|^m]^{\frac{1}{m}}$$

it always holds that

$$\sum_{i=1}^m \frac{1}{p_i} = \sum_{i=1}^m \frac{1}{m} = \frac{1}{m}\sum_{i=1}^m 1 = \frac{m}{m} = 1.$$

Comparing the SFA with the PMOO-SFA as well as analyzing the impact of dependence gives us important insight into how to calculate a performance bound for a network. As we see at the end of this chapter, our implementation enables us to conclude that the PMOO-SFA outperforms the SFA in all our examples by several orders of magnitude. Even though this does not give a proof for a general dominance of the PMOO-SFA, the sheer scale of the difference is still remarkable.

This claim differs strongly to the results of deterministic network calculus (DNC), where one can create a basic 2-chain tandem in which neither SFA nor PMOO-SFA is always better. For a proof we refer to [SZF08].

7.1. SFA with Independent Cross Flows

We assume independent cross flows. Subsequently, applying the SFA leads to this leftover service for the foi:

$$\begin{aligned}S_{\text{l.o.}} =& [S_1 - A_1]^+ \otimes [S_2 - (A_2 + A_1 \oslash S_1)]^+ \\ &\otimes \Big[S_3 - \Big(A_3 + A_2 \oslash S_2 + (A_1 \oslash S_1) \oslash [S_2 - A_2]^+\Big)\Big]^+ \otimes \\ &\cdots \otimes \Big[S_n - \Big(A_n + A_{n-1} \oslash S_{n-1} + \cdots + \Big((A_1 \oslash S_1) \oslash [S_2 - A_2]^+\Big) \oslash \\ &\cdots \oslash [S_{n-1} - (A_2 + \cdots + A_{n-1})]^+\Big)\Big]^+ .\end{aligned}$$

Remark 7.2. In the next sections, as mentioned before, we focus on the delay bounds of these sink trees. For a computation of the output bound for the SFA, see Appendix D.1.

Proposition 7.3 (Sink Tree Delay Bound with SFA)**.** *For the delay bound, we receive*

$$\begin{aligned}\mathrm{P}(d(t) > T) \overset{(2.13)}{\leq}& \mathrm{E}\Big[e^{\theta(A_{\text{foi}} \oslash S_{\text{l.o.}}\,(t+T,t))}\Big] \\ \cdots \leq& \sum_{s=0}^{t} \mathrm{E}\Big[e^{\theta A_{\text{foi}}(s,t)}\Big] \left(\sum_{s \leq l_1 \leq t+T} \cdots \sum_{l_{n-2} \leq l_{n-1} \leq t+T} \mathrm{E}\Big[e^{p_1 \theta A_1(s,l_1)}\Big]^{\frac{1}{p_1}} \mathrm{E}\Big[e^{-p_1 \theta S_1(s,l_1)}\Big]^{\frac{1}{p_1}} \right. \\ &\cdots \mathrm{E}\Big[e^{p_n \theta \left(A_n + A_{n-1} \oslash S_{n-1} + \cdots + \left(\left((A_1 \oslash S_1) \oslash [S_2 - A_2]^+\right) \oslash \cdots \oslash [S_{n-1} - (A_2 + \cdots + A_{n-1})]^+\right)\right)(l_{n-1},t+T)}\Big]^{\frac{1}{p_n}} \\ &\left. \cdot \mathrm{E}\Big[e^{-p_n \theta S_n(l_{n-1},t+T)}\Big]^{\frac{1}{p_n}} \right)\end{aligned}$$

with

$$\sum_{i=1}^{n} \frac{1}{p_i} = 1.$$

Proof. For a proof, we refer to Appendix D.2. □

Example 7.4 (Sink Tree Delay Bound with SFA for 1, 2, and 3 Servers)**.** For 1 node, this results in

$$\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} \mathrm{E}\Big[e^{\theta A_{\text{foi}}(l_0,t)}\Big] \mathrm{E}\Big[e^{\theta A_1(l_0,t+T)}\Big] \mathrm{E}\Big[e^{-\theta S_1(l_0,t+T)}\Big]$$

and for 2 nodes, this means

$$
\begin{aligned}
\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(l_0,t)}\Big] & \Bigg(\sum_{s \leq l_1 \leq t+T} \mathrm{E}\Big[e^{p_1\theta A_1(s,l_1)}\Big]^{\frac{1}{p_1}} \mathrm{E}\Big[e^{-p_1\theta S_1(s,l_1)}\Big]^{\frac{1}{p_1}} \\
& \cdot E\left[e^{p_2\theta A_2(l_1,t+T)}\right]^{\frac{1}{p_2}} E\left[e^{p_2\theta (A_1 \oslash S_1)(l_1,t+T)}\right]^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,t+T)}\right]^{\frac{1}{p_2}} \Bigg) \\
\overset{(2.21)}{\leq} \sum_{l_0=0}^{t} E\left[e^{\theta A_{foi}(l_0,t)}\right] & \Bigg(\sum_{s \leq l_1 \leq t+T} E\left[e^{p_1\theta A_1(s,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(s,l_1)}\right]^{\frac{1}{p_1}} \\
& \cdot E\left[e^{p_2\theta A_2(l_1,t+T)}\right]^{\frac{1}{p_2}} \left(\sum_{k=0}^{l_1} E\left[e^{p_2\theta A_1(k,t+T)}\right] \mathrm{E}\Big[e^{-p_2\theta S_1(k,l_1)}\Big] \right)^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,t+T)}\right]^{\frac{1}{p_2}} \Bigg),
\end{aligned}
$$

and

$$
\frac{1}{p_1} + \frac{1}{p_2} = 1,
$$

whereas for 3 nodes, our delay bound looks as follows:

$$
\begin{aligned}
& \mathrm{P}(d(t) > T) \\
& \leq \sum_{s=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(s,t)}\Big] \Bigg(\sum_{s \leq l_1 \leq t+T} \sum_{l_1 \leq l_2 \leq t+T} E\left[e^{p_1\theta A_1(s,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(s,l_1)}\right]^{\frac{1}{p_1}} \\
& \cdot E\left[e^{p_2\theta A_2(l_1,l_2)}\right]^{\frac{1}{p_2}} E\left[e^{p_2\theta (A_1 \oslash S_1)(l_1,l_2)}\right]^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,l_2)}\right]^{\frac{1}{p_2}} \\
& \cdot E\left[e^{p_3\theta A_3(l_2,t+T)}\right]^{\frac{1}{p_3}} E\left[e^{p_3\theta \left((A_1 \oslash S_1) \oslash [S_2 - A_2]^+\right)(l_2,t+T)}\right]^{\frac{1}{p_3}} E\left[e^{-p_3\theta S_3(l_2,t+T)}\right]^{\frac{1}{p_3}} \Bigg) \\
& \leq \sum_{s=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(s,t)}\Big] \Bigg(\sum_{s \leq l_1 \leq t+T} \sum_{l_1 \leq l_2 \leq t+T} E\left[e^{p_1\theta A_1(s,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(s,l_1)}\right]^{\frac{1}{p_1}} \\
& \cdot E\left[e^{p_2\theta A_2(l_1,l_2)}\right]^{\frac{1}{p_2}} \left(\sum_{k=0}^{l_1} E\left[e^{p_2\theta A_1(k,l_2)}\right] E\left[e^{-p_2\theta S_1(k,l_1)}\right] \right)^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,l_2)}\right]^{\frac{1}{p_2}} \\
& \cdot E\left[e^{p_3\theta A_3(l_2,t+T)}\right]^{\frac{1}{p_3}} \left(\sum_{u=0}^{l_2} E\left[e^{p_3\theta (A_1 \oslash S_1)(u,t+T)}\right] E\left[e^{-p_3\theta [S_2 - A_2]^+(u,l_2)}\right] \right)^{\frac{1}{p_3}} E\left[e^{-p_3\theta S_3(l_2,t+T)}\right]^{\frac{1}{p_3}} \Bigg) \\
& \overset{(2.21)}{\leq} \sum_{s=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(s,t)}\Big] \Bigg(\sum_{s \leq l_1 \leq t+T} \sum_{l_1 \leq l_2 \leq t+T} E\left[e^{p_1\theta A_1(s,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(s,l_1)}\right]^{\frac{1}{p_1}} \\
& \cdot E\left[e^{p_2\theta A_2(l_1,l_2)}\right]^{\frac{1}{p_2}} \left(\sum_{k=0}^{l_1} E\left[e^{p_2\theta A_1(k,l_2)}\right] E\left[e^{-p_2\theta S_1(k,l_1)}\right] \right)^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,l_2)}\right]^{\frac{1}{p_2}}
\end{aligned}
$$

$$\cdot E\left[e^{p_3\theta A_3(l_2,t+T)}\right]^{\frac{1}{p_3}} \left(\sum_{u=0}^{l_2}\left(\sum_{v=0}^{u} E\left[e^{p_3\theta A_1(v,t+T)}\right] E\left[e^{-p_3\theta S_1(v,u)}\right]\right) E\left[e^{p_3\theta A_2(u,l_2)}\right] E\left[e^{-p_3\theta S_2(u,l_2)}\right]\right)^{\frac{1}{p_3}}$$

$$\left.\cdot E\left[e^{-p_3\theta S_3(l_2,t+T)}\right]^{\frac{1}{p_3}}\right)$$

with

$$\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}=1.$$

Remark 7.5. We can already make some observations at this point:

1. The complexity (the "visual" as well as the computational) of the bounds increases drastically with the number of servers.
2. Since the steps to obtain the bounds are very similar and only depend on the initial description of the service, certain patterns are recognizable "pleading" for a recursion to implement the bounds.

Remark 7.6. We give bounds in the examples in particular for the three-server case since we also use this topology for some of the comparisons, e.g., between SFA and PMOO-SFA.

7.2. PMOO-SFA with Independent Cross Flow

With the same assumptions as in Section 7.1, we get for the PMOO-SFA:

$$S_{\text{l.o.}} = \left[\left(\left[\left([S_n - A_n]^+ \otimes S_{n-1}\right) - A_{n-1}\right]^+ \otimes \cdots \otimes S_1\right) - A_1\right]^+ .$$

Proposition 7.7 (Sink Tree Delay Bound with PMOO-SFA)**.** *For the delay bound, we obtain*

$$\mathrm{P}(d(t)>T) \overset{(2.13)}{\leq} \mathrm{E}\Big[e^{\theta A_{\text{foi}}\oslash S_{\text{l.o.}}(t+T,t)}\Big]$$

$$\cdots \leq \sum_{s=0}^{t} \mathrm{E}\Big[e^{\theta A_{\text{foi}}(s,t)}\Big]\,\mathrm{E}\Big[e^{\theta A_1(s,t+T)}\Big]\left(\sum_{l_1=s}^{t+T}\mathrm{E}\Big[e^{-\theta S_1(l_1,t+T)}\Big]\,\mathrm{E}\Big[e^{\theta A_2(s,l_1)}\Big]\right.$$

$$\cdot\left(\sum_{l_2=s}^{l_1}\mathrm{E}\Big[e^{-\theta S_2(l_2,l_1)}\Big]\,\mathrm{E}\Big[e^{\theta A_3(s,l_2)}\Big]\cdots\left(\sum_{l_k=s}^{l_{k-1}}\mathrm{E}\Big[e^{-\theta S_k(l_k,l_{k-1})}\Big]\,\mathrm{E}\Big[e^{\theta A_{k+1}(s,l_k)}\Big]\right.\right.$$

$$\left.\left.\left.\cdots\left(\sum_{l_{n-1}=s}^{l_{n-2}}\mathrm{E}\Big[e^{-\theta S_{n-1}(l_{n-1},l_{n-2})}\Big]\,\mathrm{E}\Big[e^{\theta A_n(s,l_{n-1})}\Big]\,\mathrm{E}\Big[e^{-\theta S_n(s,l_{n-1})}\Big]\right)\right)\right)\right).$$

Proof. See Appendix D.3. □

Remark 7.8. We observe that the bound could be calculated without applying Hölder's inequality. This is why we conclude that the dependency observed for SF-Analysis is not given by the system but rather by the SFA itself. We come back to this point in our comparison in Section 7.3.

Example 7.9 (Sink tree delay bound with PMOO-SFA for 1, 2 and 3 servers and independent cross traffic)**.** For 1 node, this (again) results in

$$\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(l_0,t)}\Big] \,\mathrm{E}\Big[e^{\theta A_1(l_0,t+T)}\Big] \,\mathrm{E}\Big[e^{-\theta S_1(l_0,t+T)}\Big],$$

but for 2 nodes, we get

$$\begin{aligned}\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(l_0,t)}\Big] \,\mathrm{E}\Big[e^{\theta A_1(l_0,t+T)}\Big] &\left(\sum_{l_1=l_0}^{t+T} \mathrm{E}\Big[e^{-\theta S_1(l_1,t+T)}\Big] \,\mathrm{E}\Big[e^{\theta A_2(l_0,l_1)}\Big]\right.\\ &\left.\cdot \,\mathrm{E}\Big[e^{-\theta S_2(l_0,l_1)}\Big]\right).\end{aligned}$$

The example for 3 nodes yields

$$\begin{aligned}\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}}(l_0,t)}\Big] \,\mathrm{E}\Big[e^{\theta A_1(l_0,t+T)}\Big] &\left(\sum_{l_1=l_0}^{t+T} \mathrm{E}\Big[e^{-\theta S_1(l_1,t+T)}\Big] \,\mathrm{E}\Big[e^{\theta A_2(l_0,l_1)}\Big]\right.\\ &\left.\cdot \left(\sum_{l_2=l_0}^{l_1} \mathrm{E}\Big[e^{-\theta S_2(l_2,l_1)}\Big] \,\mathrm{E}\Big[e^{\theta A_3(l_0,l_2)}\Big] \,\mathrm{E}\Big[e^{-\theta S_3(l_0,l_2)}\Big]\right)\right).\end{aligned}$$

7.3. SFA / PMOO-SFA Comparison

We have implemented the delay bound for the three-server case. The results are depicted in a log-plot in Figure 7.2. We drew an auxiliary line at the y-value equal to 1, since our delay bound limits a probability from above and thus, each value > 1 is actually not a useful result. By "optimized", we mean, as mentioned in the introduction of Chapter 7, that we performed a systematic "grid search". If θ is not optimized, it is set equal to 1.

We see that in this setting, the PMOO outperforms the SFA drastically by several orders of magnitude and that parameter optimization is a huge factor. For example, for $T = 8$, the delay bound for the SFA is equal to $1.12 \cdot 10^5$ whereas the PMOO-SFA leads to a bound equal to $1.3 \cdot 10^{-4}$. By parameter optimization, the SFA improves at $T = 8$ to 349.8 (still way above 1) and the PMOO-SFA drops down to $1.3{\cdot}10^{-10}$. We further investigate the issue of parameter optimization in Chapter 8.

As an overall conclusion, we notice that the SFA provides highly undesirable bounds caused by its "artificial dependencies". By artificial dependency we mean, in this context, that the dependent factors inside the expectation could have been avoided by using a different type of analysis. As a consequence, we continue our analysis only with the PMOO-SFA hereafter.

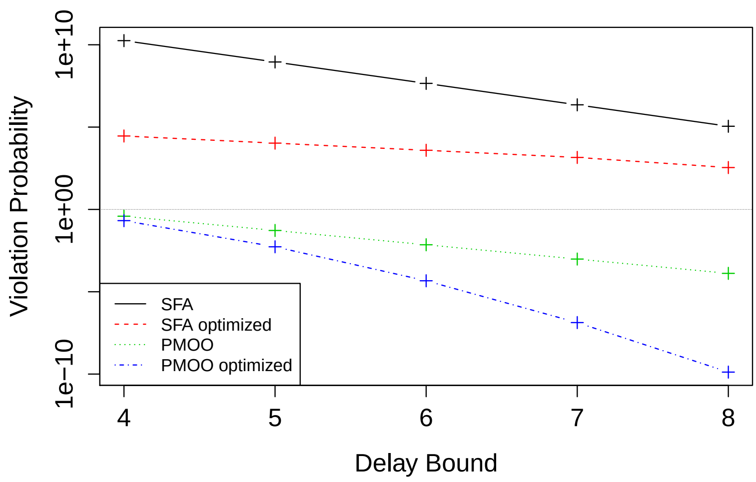


Figure 7.2.: Delay bound comparison of the SFA and the PMOO-SFA in a sink tree for $t = 20$, $n = 1$, $\rho = 1$, and $r = 5$.

7.4. PMOO-SFA with Dependent Cross Flow

Figure 7.2 gave us a first impression on how dependencies can influence the resulting bounds. In this subsection, we compute the delay bound via the PMOO-SFA for dependent cross flows and compare it to the bound obtained in Subsection 7.2. Hence, this comparison is well-suitable to determine the effect of dependence and its consequence of applying Hölder's inequality.

Proposition 7.10 (Sink tree delay Bound with PMOO-SFA and dependent cross-flows). *We obtain for the delay bound*

$$
\begin{aligned}
\mathrm{P}(d(t) > T) &\overset{(2.13)}{\leq} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o.}}(t+T,t)}\right] \\
\cdots &\leq \sum_{s=0}^{t} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] \left(\mathrm{E}\left[e^{p_1 \theta A_1(s,t+T)}\right]\right)^{\frac{1}{p_1}} \\
&\quad \cdot \left(\sum_{l_1=s}^{t+T} \mathrm{E}\left[e^{-p_2 \theta S_1(l_1,t+T)}\right] \left(\mathrm{E}\left[e^{p_2 p_3 \theta A_2(s,l_1)}\right]\right)^{\frac{1}{p_3}} \right. \\
&\quad \cdots \left(\sum_{l_{k-1}=s}^{l_{k-2}} E\left[e^{-p_2 p_4 \cdots p_{2k-2} \theta S_{k-1}(l_{k-1},l_{k-2})}\right] \left(E\left[e^{p_2 p_4 \cdots p_{2k-2} p_{2k-1} \theta A_k(s,l_{k-1})}\right]\right)^{\frac{1}{p_{2k-1}}} \right. \\
&\quad \cdots \left(\sum_{l_{n-2}=s}^{l_{n-3}} E\left[e^{-p_2 p_4 \cdots p_{2n-4} \theta S_{n-2}(l_{n-2},l_{n-3})}\right] \left(E\left[e^{p_2 p_4 \cdots p_{2n-4} p_{2n-3} \theta A_{n-1}(s,l_{n-1})}\right]\right)^{\frac{1}{p_{2n-3}}} \right.
\end{aligned}
$$

$$
\cdot \left(\sum_{l_{n-1}=s}^{l_{n-2}} E\left[e^{-p_2 p_4 \cdots p_{2n-2}\theta S_{n-1}(l_{n-1},l_{n-2})}\right] E\left[e^{p_2 p_4 \cdots p_{2n-2}\theta A_n(s,l_{n-1})}\right] \right.
$$

$$
\left. \cdot\, E\left[e^{-p_2 p_4 \cdots p_{2n-2}\theta S_n(s,l_{n-1})}\right] \right)^{\frac{1}{p_{2n-2}}}
$$

$$
\left.\left.\left. \ldots \right)^{\frac{1}{p_{2k-2}}} \right)^{\frac{1}{p_4}} \right)^{\frac{1}{p_2}}
$$

with

$$
\begin{aligned}
\frac{1}{p_1} + \frac{1}{p_2} &= 1 \\
\vdots &= \vdots \\
\frac{1}{p_{2n-1}} + \frac{1}{p_{2n-2}} &= 1.
\end{aligned}
$$

Proof. The proof contains a quite involved computation (see Appendix D.4). □

Example 7.11 (Sink tree delay bound with PMOO-SFA for 1, 2 and 3 servers and dependent cross traffic)**.** For 1 server, this results again (due to a lack of dependent factors) in

$$
\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}}(l_0,t)}\right] \mathrm{E}\left[e^{\theta A_1(l_0,t+T)}\right] \mathrm{E}\left[e^{-\theta S_1(l_0,t+T)}\right],
$$

but for 2 servers, we obtain

$$
\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(l_0,t)}\right] E\left[e^{p_1\theta A_1(l_0,t+T)}\right]^{\frac{1}{p_1}} \left(\sum_{l_1=l_0}^{t+T} E\left[e^{-p_2\theta S_1(l_1,t+T)}\right] E\left[e^{p_2\theta A_2(l_0,l_1)}\right] \right.
$$

$$
\left. \cdot E\left[e^{-\theta S_2(l_0,l_1)}\right] \right)^{\frac{1}{p_2}}
$$

and

$$
\frac{1}{p_1} + \frac{1}{p_2} = 1.
$$

Applying this formula to our 3 servers example gives us:

$$
\begin{aligned}
\mathrm{P}(d(t) > T) \leq \sum_{l_0=0}^{t} & E\left[e^{\theta A_{\mathrm{foi}}(l_0,t)}\right] E\left[e^{p_1 \theta A_1(l_0,t+T)}\right]^{\frac{1}{p_1}} \\
& \cdot \left(\sum_{l_1=l_0}^{t+T} E\left[e^{-p_2 \theta S_1(l_1,t+T)}\right] \left(E\left[e^{p_2 p_3 \theta A_2(l_0,l_1)}\right]\right)^{\frac{1}{p_3}} \right. \\
& \left. \cdot \left(\sum_{l_2=l_0}^{l_1} E\left[e^{-p_2 p_4 \theta S_2(l_2,l_1)}\right] E\left[e^{p_2 p_4 \theta A_3(l_0,l_2)}\right] E\left[e^{-p_2 p_4 \theta S_3(l_0,l_2)}\right] \right)^{\frac{1}{p_4}} \right)^{\frac{1}{p_2}}
\end{aligned}
$$

with

$$
\begin{aligned}
\frac{1}{p_1} + \frac{1}{p_2} &= 1 \\
\frac{1}{p_3} + \frac{1}{p_4} &= 1.
\end{aligned}
$$

7.5. Dependence / Independence Comparison

We implemented the delay bounds obtained by the PMOO-SFA analysis for independent and dependent cross traffic. The results are shown in a log-plot in Figure 7.3. Again, the parameters are given at the top of the graphic. The observation meets our expectation quite well: again, the dependent case in which the Hölder inequality is applied leads to much worse results. For the standard parameter choice, i.e., $\theta = 1$ and $p_i = 2$, the bound for dependent cross flow decreases from 15.5 to 0.005 as T goes from 4 to 8 whereas for the independent case it decreases from $5.2 \cdot 10^{-3}$ to $3.2 \cdot 10^{-8}$. The optimized version of the independent case performs much better with a delay bound of $1.7 \cdot 10^{-24}$ at $T = 8$.

We conclude that the application of Hölder's inequality is highly unfavorable for the tightness of our analysis due to its gap of several orders of magnitude. It even suggests that we should minimize the application of the Hölder inequality with the highest priority to avoid its severe side effects. In particular, for the independent case we also see that optimization is able to improve the bound tremendously, leading us to investigate this issue in the next chapter.

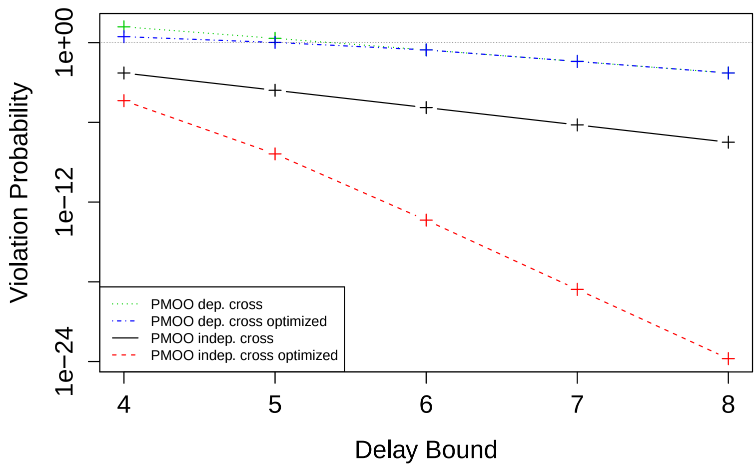


Figure 7.3.: Delay bound comparison of the PMOO-SFA for dep./indep. cross traffic in a sink tree for $t = 20$, $n = 1$, $\rho = 1$, and $r = 6$.

8. Optimizing the Parameters with Derivative-free Methods

In the Subsections 7.3 and 7.5, we have already seen that optimizing θ as well as the Hölder parameters is crucial for our analysis since the deviation from a standard choice for the parameters is very high (this was observed in the Figures 7.2 and 7.3). In case we only have to optimize the MGF parameter θ, at a first sight, the delay bound obtained in Subsection 7.2 gives us a presumably convex shape, as we see in Figure 8.1a. Unfortunately, as we prove now, the bound is not convex in general.

But beforehand, let us condense the results of this chapter. At first, we prove for the looser bound (8.2), that it is always convex, but not the delay bound using this MGF-bound. This and a complicated first derivative influenced us to focus on derivative-free optimization, i.e., in the following we only apply optimization algorithms, where the knowledge (actually even the existence) of the derivative is not necessary.

Then, we present four different (derivative-free) optimization algorithms and some of their properties. At the end of this chapter, we compare optimizations' results of our implementation. This is useful at two levels:

1. We are able to show the advantages of some of the various algorithms (especially the pattern search and simulated annealing appeared quite favorable due to their better performance).

2. We gain information about the shape of the bound. We show below that, especially for many servers, a convex shape is rather unlikely.

Now, we prove that the MGF-bound (first part), but not the delay bound (second part), is always convex. So at first, let us restate which bound is used. For the MGF-bound we obtained by applying the Conversion Theorem (Theorem 2.9) to the tail-bound in [MB00] for adversarial traffic, we received eventually

$$
\begin{aligned}
\mathrm{E}\left[e^{\theta A(s,t)}\right] &\overset{(2.4)}{\leq} e^{\theta\rho(t-s)} \int_0^1 e^{\theta\sqrt{-\log(\varepsilon)2\sum_{i=1}^n \sigma_i^2}} \mathrm{d}\varepsilon \\
&\overset{(4.9)}{\leq} e^{\theta\rho\cdot(t-s)} \left(1 + \sqrt{\frac{\pi}{2}\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n \sigma_i^2} \left(\mathrm{erf}\left(\theta\sqrt{\frac{1}{2}\sum_{i=1}^n \sigma_i^2}\right) + 1\right)\right) \quad (8.1) \\
&\overset{(4.10)}{\leq} e^{\theta\rho\cdot(t-s)} \left(1 + \sqrt{2\pi\sum_{i=1}^n \sigma_i^2}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^n \sigma_i^2}\right) =: f(\theta), \quad (8.2)
\end{aligned}
$$

where erf(.) denotes the Gauss error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathrm{d}t.$$

Due to our constant service assumption, we get for the MGF of the service that

$$\mathrm{E}\left[e^{-\theta S(s,t)}\right] = e^{-\theta r(t-s)}. \tag{8.3}$$

Proposition 8.1. *The MGF-Bound in (8.2) is convex in θ.*

Proof. We derive (8.2) twice for θ and obtain via Wolfram Alpha[1]:

$$\begin{aligned}
\frac{\partial^2 f(\theta)}{\partial \theta^2} =& \frac{\partial^2}{\partial \theta^2} e^{\theta\rho\cdot(t-s)} \left(1 + \sqrt{2\pi \sum_{i=1}^n \sigma_i^2 \theta e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2}}\right) \\
\approx& \underbrace{\rho^2 (t-s)^2 \left(\theta \sqrt{2\pi \sum_{i=1}^n \sigma_i^2 e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2}} + 1\right) e^{\rho\theta(t-s)}}_{>0} \\
&+ \underbrace{2\rho(t-s) \left(2.50663 \sigma^{\frac{3}{2}} \theta^2 e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2} + \sqrt{2\pi \sum_{i=1}^n \sigma_i^2 e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2}}\right) e^{\rho\theta(t-s)}}_{>0} \\
&+ \left(\underbrace{5.01326 \sigma^{\frac{3}{2}} \theta e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2} + \theta \sqrt{2\pi \sum_{i=1}^n \sigma_i^2 e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2}}}_{>0} \right. \\
&\left. \underbrace{\cdot \left(\left(\theta \sum_{i=1}^n \sigma_i^2\right)^2 + \sum_{i=1}^n \sigma_i^2 \right) \right) e^{\rho\theta(t-s)}}_{>0},
\end{aligned}$$

which proves that the looser MGF-bound (8.2) is convex in θ. $\square$

Proposition 8.2. *The delay bound for the one-server case with MGF-Bound (8.2)*

$$\begin{aligned}
\mathrm{P}(d(t) > T) &\leq \sum_{l_0=0}^{t} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}}(l_0,t)}\right] \mathrm{E}\left[e^{\theta A_1(l_0,t+T)}\right] \mathrm{E}\left[e^{-\theta S_1(l_0,t+T)}\right] \\
&\overset{(8.2)(8.3)}{\leq} \sum_{l_0=0}^{t} e^{\theta\rho(t-l_0)} \left(1 + \sqrt{2\pi \sum_{i=1}^n \sigma_i^2 \theta e^{\frac{1}{2}\theta^2 \sum_{i=1}^n \sigma_i^2}}\right)
\end{aligned}$$

[1]https://www.wolframalpha.com/

$$\cdot\, e^{\theta\rho(t+T-l_0)}\left(1+\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}\right)e^{-\theta r(t+T-l_0)}$$

$$=\sum_{l_0=0}^{t}\left(1+\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}\right)^2 e^{\theta(\rho(2t+T-2l_0)-r(t+T-l_0))} =: \sum_{l_0=0}^{t} g_{l_0}(\theta)$$

is not convex in general.

Proof. In order to prove that the delay bound is not convex in general, we only have to give an example where this does not hold true. But before, we make some preliminary considerations. Let us define $c_0 := \rho\,(2t+T-2l_0) - r\,(t+T-l_0)$ and thus, the second derivative for a single summand yields according to Wolfram Alpha:

$$\begin{aligned}
\frac{\partial^2 g_{l_0}(\theta)}{\partial\theta^2} =& \underbrace{c_0^2 e^{c_0\theta}\left(\theta\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}+1\right)^2}_{>0} \\
&+ \underbrace{4c_0}_{\gtrless 0}\, \underbrace{e^{c_0\theta}\left(2.50663\left(\sum_{i=1}^{n}\sigma_i^2\right)^{\frac{3}{2}}\theta^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}+\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}\right)}_{>0} \\
&\cdot \underbrace{\left(\theta\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}+1\right)}_{>0} \\
&+ \underbrace{e^{c_0\theta}}_{>0}\Bigg(\underbrace{2\left(2.50663\left(\sum_{i=1}^{n}\sigma_i^2\right)^{\frac{3}{2}}\theta^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}+\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}\right)^2}_{>0} \\
&+ \underbrace{2\left(\theta\sqrt{2\pi\sum_{i=1}^{n}\sigma_i^2 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}}+1\right)}_{>0} \\
&\cdot \underbrace{\left(7.51988\left(\sum_{i=1}^{n}\sigma_i^2\right)^{\frac{3}{2}}\theta e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}+2.50663\left(\sum_{i=1}^{n}\sigma_i^2\right)^{\frac{5}{2}}\theta^3 e^{\frac{1}{2}\theta^2\sum_{i=1}^{n}\sigma_i^2}\right)}_{>0}\Bigg),
\end{aligned}$$

which means that the convexity of the delay bound depends on the sign of $c_0 = \rho\,(2t+T-2l_0) - r\,(t+T-l_0)$. If c_0 is positive, then the bound is convex, but if c_0 is negative then we could conclude at this point neither concavity nor convexity without computing the derivative explicitly. See Example 8.4 where we have actually calculated the bound with the proclaimed negative sign. This finishes the proof. $\square$

Remark 8.3. We would like to mention that since we assume the utilization to be smaller than 1, we have $r > \rho$ and thus c_0 is very likely to be negative.

Example 8.4. For the parameters $s = 1$, $t = 1$, $T = 3$, $\theta = 1$, $\sum_{i=1}^{n} \sigma_i^2 = 1$, $\rho = 1$, $n = 1$, and $r = 10$ we calculate for the second derivative $-1.19 \cdot 10^{-9} < 0$, so the delay bound is not convex (even though this value is very close to 0).

Remark 8.5. The bound in general is also strongly influenced by the Hölder parameter (Figure 8.1b Figure 8.1c).

For instance, if we apply the classic Hölder inequality to two factors, we receive a $p \geq 1$ and a $q \geq 1$ that depends on p by $\frac{1}{p} + \frac{1}{q} = 1 \Rightarrow q = \frac{1}{1-\frac{1}{p}} = \frac{1}{\frac{p-1}{p}} = \frac{p}{p-1}$. By looking at the relation of p and q in Figure 8.2 we conclude that the parameter sensitivity is quite high for $p < 2$, since $q'(2) = -1$, $q'\left(\frac{3}{2}\right) = -4$ and it holds true that

$$\lim_{p \searrow 1} \frac{\partial q(p)}{\partial p} = \lim_{p \searrow 1} -\frac{1}{(p-1)^2} = -\infty.$$

Since we know that at least one of the parameters is always ≤ 2, a certain level of parameter sensitivity is always given.

We compare different ways of optimizing parameters by the bound they provide in relation to the number of evaluations of the bound, so-called “function calls”. We obtain terms whose complexity “explodes” with the number of servers which leads to the fact that it takes a long time to evaluate the function. It is the bottleneck of the whole optimization and hence, we would like to limit this by upper bounding the number of function calls.

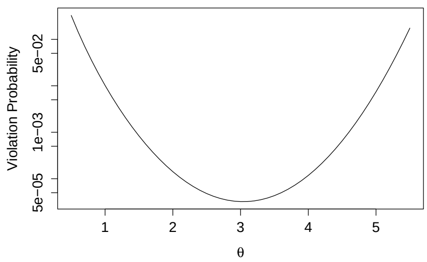
In the following sections, we demonstrate several methods. They all have in common that we do not compute a derivative, so we only have to assume our objective function to be continuous. We have implemented the bound in the statistical programming language **R**[2], version 3.3.2, and in the general-purpose programming language **Python**[3], version 3.5.2. The network setting which gives us the parameters to be optimized is a sink tree with dependent cross flow after applying the PMOO analysis (Subsection 7.4), i.e., we optimize θ and the Hölder parameters. Concerning the implementation, we remark that we do it by ourselves step-by-step instead of relying on presumably faster packages that contain the algorithms. Hence, we can maintain a certain control of the computation.

8.1. Systematic Search

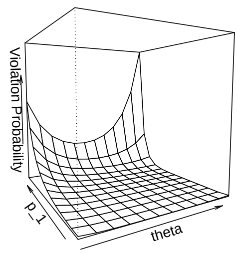
In systematic search, the idea is to specify a set of parameters in which we search the combination that minimizes the objective. More precisely, we define a set I s.t.

$$\min_{x_i \in I} f(x_i)$$

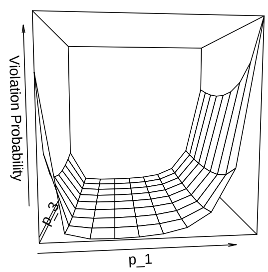
[2]https://www.r-project.org
[3]https://www.python.org



(a) Delay bound for 2 servers (indep. cross flows).



(b) Delay bound for 2 servers according to PMOO-SFA (dep. cross flows).



(c) Delay bound for 3 servers according to PMOO-SFA (dep. cross flows, θ is set to 1).

Figure 8.1.: Parameter sensitivity of the delay bound.

and compute the minimum by trying all x_i. Of course, in order to be computationally feasible, we have to ensure that I is not too “big” which is not always straightforward, especially since we have to cope with multidimensional functions.

In order to provide a finite computation time, the parameter set are discrete points along the grid and bounded, even though the parameter space itself is only lower bounded. Otherwise, the systematic search could not have been executed in finite time at all. We limit the parameter space for θ most of the times to the set $[0.5, 3.5]$ and for each Hölder parameter to $[1.5, 3.5]$, leading to $\left[\frac{3.5}{3.5-1}, \frac{1.5}{1.5-1}\right] = [1.4, 3]$ for the other part of each Hölder pair. The fact that the optimal values for the other optimization algorithms lie in this interval shows that we did not place the systematic search at a disadvantage; conversely, it benefits from the knowledge obtained by the other algorithms. A reason why such a small space is possibly already sufficient to find the global optimum (high parameter sensitivity) is given in the beginning of this chapter as well as in Figure 8.1.

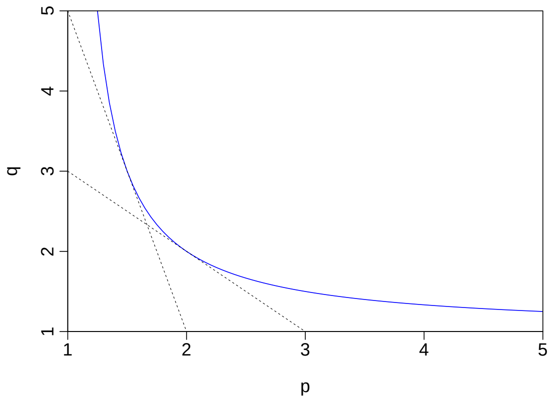


Figure 8.2.: Relation of the Hölder parameters: $q(p) = \frac{p}{p-1}$.

8.2. Pattern Search

This special case of "Direct Search" ([HJ61]) is a simple way to minimize a function. We define a "base point" which is an arbitrary feasible solution and the starting point of the algorithm. Then we continue with an "Exploratory move" which changes each parameter by a certain predetermined δ and checks whether an improvement is obtained or not. After running the exploratory move through each parameter, we define a new base point. In case we were able to find a new base point that has a lower objective, this step is followed by a so-called "Pattern move", i.e., we change each parameter in the direction of the new base point. If this step does not give an improvement or if we even fail in the explanatory move, we decrease the δ. The algorithm terminates as soon as a target δ is reached or in our case, the maximum number of function calls is reached.

8.3. Nelder–Mead Method

This algorithm was proposed in [NM65]. Given a function $f : X \to \mathbb{R}$ to be minimized with $X \subseteq \mathbb{R}^n$, we start with a simplex, i.e., a set consisting of $n+1$ points. We then use operations called "reflection" or "contraction" which basically compute a new point in each iteration that is on a line with the worst point (which in this context means the point with the largest function value) and the simplex's centroid. If this new point gives us an improvement, the algorithm replaces the worst point with the new point in the current "simplex set". Simulations exhibited that the result heavily depends on the starting simplex, i.e., it is not capable of overcoming a local optimum easily. On the other hand, convergence is quite fast as we see in Section 8.5.

8.4. Simulated Annealing (SA)

Simulated annealing ([KGV83]) is a completely different approach than the aforementioned analyses. It evaluates randomly chosen points in the neighborhood and also accepts non-improving steps ("uphill climbs") with a certain probability. This procedure enables it to overcome local optima on its way to a global one. The idea is to simulate a cooling scheme observed in statistical mechanics, that means that we allow for coarse changes in our search when the "temperature" (in our case a real number) is high and decrease the acceptance probability as the temperature slowly "cools down". The connection of temperature and acceptance of a change in our search variable is provided by the so-called Metropolis rule (for further details, see again [KGV83]).

8.5. Evaluation

As we mentioned in the beginning of this chapter, we evaluate the different algorithms by the following metric: We limit the possible number of function calls, interrupt the algorithm if the limit is reached and look at the current optimal solution, i.e., we look at a sort of "evaluation efficiency". Our benchmark is the delay bound with the standard choice for the parameters without any optimization. Therefore, we chose $p_i = 2$ and θ according to the table below.

We calculated the delay bounds for different numbers of servers in a sink tree that has been analyzed with the PMOO-SFA assuming the cross flows to be dependent; the plots are depicted in Figure 8.3 and 8.4. We were also able to calculate the delay bounds for an extended number of servers (5 or more), but mostly omitted the systematic search in this case since the number of function calls and thus the execution times explode (see Figure 8.5). For the parameters, we chose $t = 20, T = 5, \sum_{i=1}^{n} \rho_i = 1$ and $\sum_{i=1}^{n} \sigma_i = 1$ with $\rho_i = \sigma_i = \frac{1}{n}$ and $n = 1$. In order to provide a computationally feasible result, θ (if not optimized) and r are chosen according to the following table:

Number of servers	θ	r
2	1	4.5
3	0.5	6
4	0.5	10
5	0.5	12
10	0.01	22

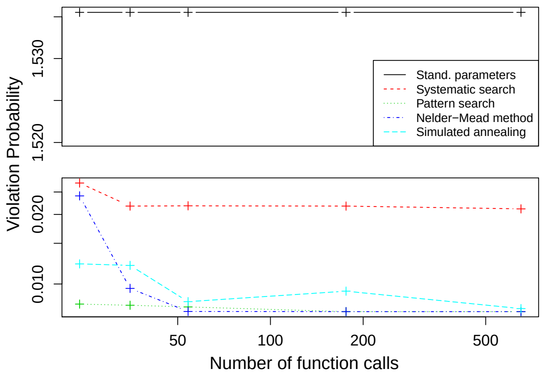
Every value on the x-axis is not a continued run but a completely new one. This explains why for two algorithms, the delay bound sometimes increases even though the number of function calls increases, especially for a randomized algorithm as simulated annealing. The second algorithm prone to this issue is the systematic search in case the new (smaller) step size is not a divisor of the previous one. For a more detailed explanation, see the next example.

The difference between some standard choices and parameter optimization is tremendous, regardless of the number of servers, even though we give the standard bound an advantage by decreasing θ to improve its function value (e.g., for 10 servers, the maximal solution that does not exceed the maximal number possible in R and in Python is below 0.1). This is the reason why we have split the plot at the y-axis.

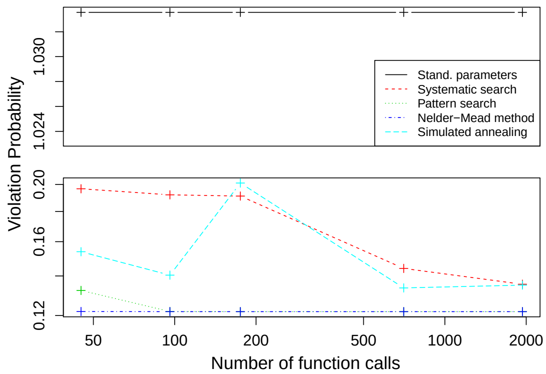
Comparing the optimization techniques themselves up to 5 servers we see that, in terms of “evaluation efficiency”, the systematic search performs quite badly, even though it is, in theory, capable of overcoming local optima. This might indicate that in practice, non-convexity is not a real issue for this topology but one should rather focus on increased precision. Simulated annealing, due to its in-built variability, has a big variance and does not perform well, either. According to our plots, the top performers are the Nelder-Mead method and the pattern search. It took more time to fine-tune Nelder-Mead, since the initial simplex has to be set carefully, but if this is accomplished, the obtained bounds are very compatible to pattern search outperforming the other methods significantly with quick convergence and stable results.

If we look at the 10-server case (Figure 8.5c), the comparison leads to quite different results, in particular the fact that simulated annealing outperforms the other algorithms drastically. This can be interpreted in a way that for more complicated settings, no convexity property can be assumed anymore since the only algorithm capable of uphill climbing obtains a significantly lower delay bound. Another advantage of simulated annealing is that it depends less on the initial value due to its randomness. On the contrary, the other algorithms get “stuck” and only manage to improve hardly over time.

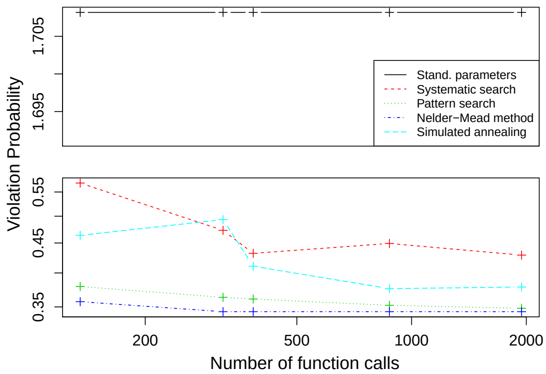
Example 8.6 (Smaller δ can lead to worse bounds for the systematic search)**.** Let us assume that θ is the only parameter to be optimized, that the optimum is unique and attained at $\theta = 1$. We optimize θ by searching between 0.5 and 3.5. If $\delta = 0.5$, i.e., the search space is $0.5, 1, \ldots 3.5$ and we find the optimal value. Decreasing δ to 0.4 changes the search space to $0.5, 0.9, 1.3, \ldots 3.3$ and thus the result gets worse even though the granularity is increased. This flaw can only be prevented by ensuring that each new δ is a divisor of the old one. But in this case, the search space grows exponentially in each step.



(a) Sink tree with two servers.

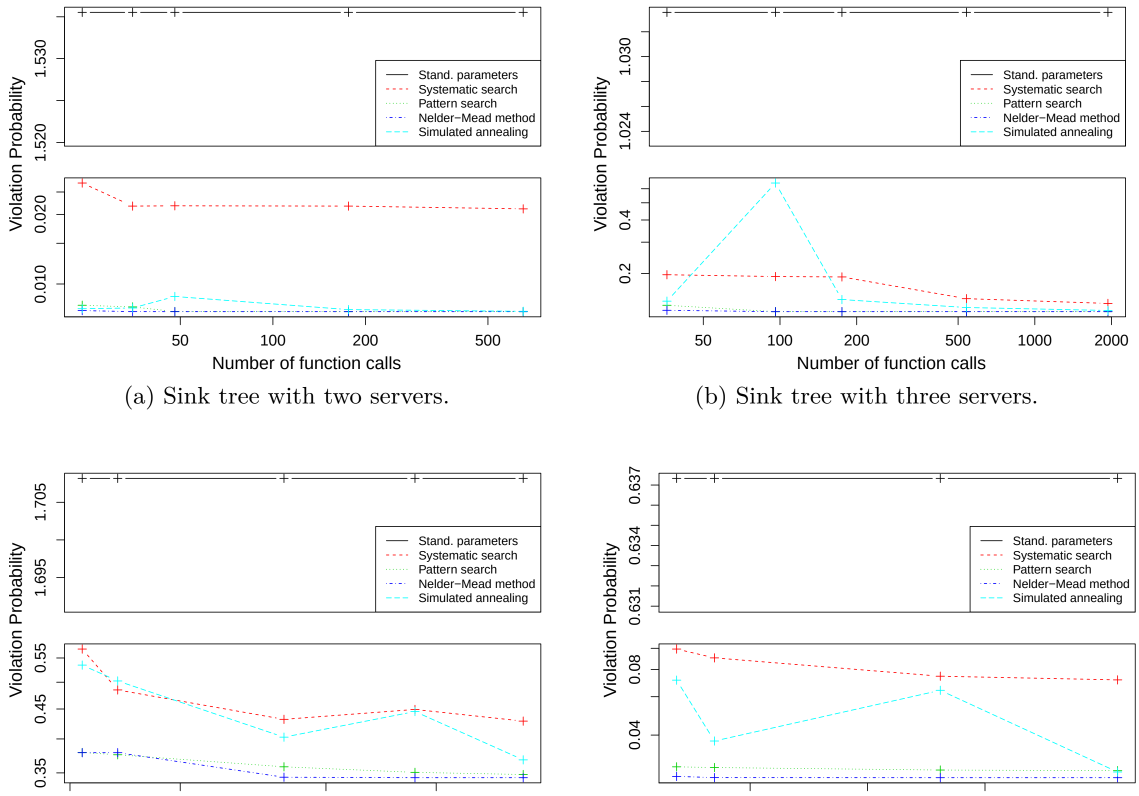


(b) Sink tree with three servers.



(c) Sink tree with four servers.

Figure 8.3.: Comparison of the delay bound for dep. cross traffic in a sink tree for different optimization algorithms (Python).



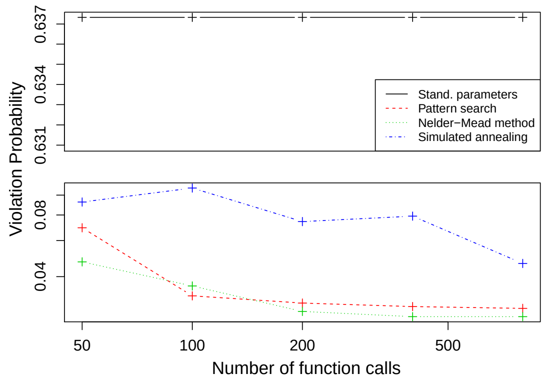
(a) Sink tree with two servers.

(b) Sink tree with three servers.

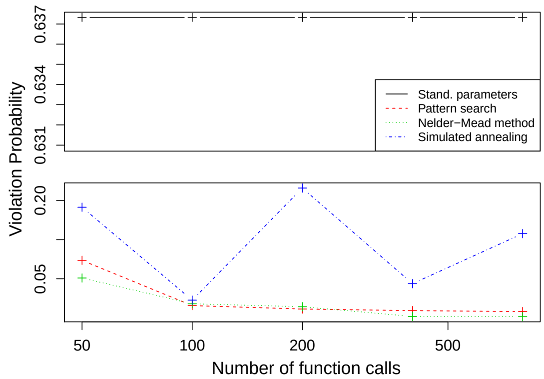
(c) Sink tree with four servers.

(d) Sink tree with five servers.

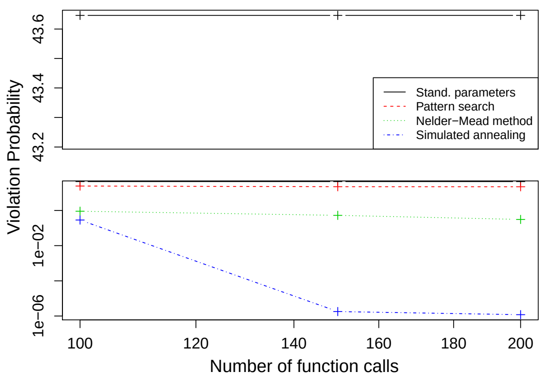
Figure 8.4.: Comparison of the delay bound for dep. cross traffic in a sink tree for different optimization algorithms (R).



(a) Sink tree with five servers (Python).



(b) Sink tree with five servers (R).



(c) Sink tree with ten servers (R).

Figure 8.5.: Comparison of the delay bound for dep. cross traffic in a sink tree without the systematic search.

Part IV.

Improving the Output Bound with Lyapunov's Inequality

9. Obtaining a Better Bound by Inserting an Additional Inequality

9.1. Motivation of an Adjusted Version of the Standard Approach

Stochastic Network Calculus and its MGF-bound framework (see for example [Fid06]) provides us with elegant tools, but the tightness of these bounds is seen as rather questionable. Work as in [CPS13] even argue that this idea, which we we call "standard approach" in the following, is doomed from the beginning and should be replaced by martingale bounds. To the best of our knowledge, no paper on a network analysis including multi-hops with martingale bounds has been published so far. This gives us a strong impetus to try to improve the standard approach since it has the advantage of still being able to use the network analysis properties.

In this part, we modify the standard approach to compute the output bound,

$$\begin{aligned}\phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\leq} \mathrm{E}\Big[e^{\theta A \oslash S\,(s,t)}\Big] \\ &\overset{(2.7)}{=} \mathrm{E}\Big[e^{\theta \max_{0\leq i\leq s}\{A(i,t)-S(i,s)\}}\Big] \\ &= \mathrm{E}\Big[\max_{0\leq i\leq s} e^{\theta(A(i,t)-S(i,s))}\Big] \\ &\overset{(2.21)}{\leq} \sum_{i=0}^{s} \mathrm{E}\Big[e^{\theta(A(i,t)-S(i,s))}\Big],\end{aligned} \tag{9.1}$$

and show that if we insert an additional bound in this chain of inequalities, then this new bound is always at least as good as the standard bound. Since we only change the derivation slightly, all the network analysis properties are still intact. After laying the foundation in this chapter, we implement output bounds for several different distributions in Chapter 10 and show that this new approach leads, in most cases, to better bounds.

9.2. Introduction of the "Lyapunov Bound"

Inequality (2.21), we called "max-sum inequality" in Section 2.5, is on the one hand, simple and intuitive but on the other hand, very coarse as this easy example uncovers:

Let $x_1, \dots, x_t \in \mathbb{R}$ with $x_i = x_1$ for all $i = 1, \dots, t$. In this case, the max-sum inequality yields

$$\begin{aligned} e^{x_1} &= e^{\max_{1 \le i \le t} x_i} \\ &\overset{(2.21)}{\le} \sum_{i=0}^{t-1} e^{x_i} \end{aligned}$$

and we already see that the more similar the x_i are, the bigger the gap becomes. This step, which is equivalent to the “Union Bound” for the calculation of performance bounds (see Appendix E), has been strongly criticized by past literature ([CPS13]). We show now that it is a possible attempt to “tighten up” the max-sum inequality by applying Lyapunov's inequality in between.

Theorem 9.1 (Lyapunov's Inequality)**.** *Let X be a random variable. For $1 \le p < \infty$, it holds true that*

$$\mathrm{E}[X] \le \mathrm{E}[|X|] \le \left(\mathrm{E}[|X|^p]\right)^{\frac{1}{p}}, \tag{9.2}$$

and that

$$\mathrm{E}[|X|] = \inf_{p \ge 1} \left(\mathrm{E}[|X|^p]\right)^{\frac{1}{p}}. \tag{9.3}$$

Proof. The first inequality follows by the monotonicity of the expectation. The second follows by Jensen's inequality: It can easily be shown that the p-th root function is concave for $p \ge 1$ (see Proposition F.1). By Jensen's inequality ((F.2) in the Appendix), it holds

$$\mathrm{E}[|X|] = \mathrm{E}\Big[\left(|X|^p\right)^{\frac{1}{p}}\Big] \overset{(F.2)}{\le} \left(\mathrm{E}[|X|^p]\right)^{\frac{1}{p}}.$$

The second claim is trivial since $p = 1$ solves (9.2) with equality. □

With Theorem 9.1, we can extend the standard approach (9.1) by our new output bound:

Theorem 9.2. *It holds true that*

$$\phi_{A'(s,t)}(\theta) \le \inf_{p \ge 1} \left\{ \left(\sum_{i=0}^{s} \mathrm{E}\Big[e^{p\theta(A(i,t) - S(i,s))}\Big] \right)^{\frac{1}{p}} \right\}. \tag{9.4}$$

Proof. We compute

$$\begin{aligned} \phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\le} \mathrm{E}\Big[e^{\theta A \oslash S\,(s,t)}\Big] \\ &\overset{(2.7)}{=} \mathrm{E}\Big[e^{\theta \max_{0 \le i \le s}\{A(i,t) - S(i,s)\}}\Big] \\ &\overset{(9.3)}{=} \inf_{p \ge 1} \left\{ \left(\mathrm{E}\Big[e^{p\theta \max_{0 \le i \le s}\{A(i,t) - S(i,s)\}}\Big] \right)^{\frac{1}{p}} \right\} \\ &\overset{(2.21)}{\le} \inf_{p \ge 1} \left\{ \left(\sum_{i=0}^{s} \mathrm{E}\Big[e^{p\theta(A(i,t) - S(i,s))}\Big] \right)^{\frac{1}{p}} \right\}. \end{aligned} \tag{9.5}$$

□

Remark 9.3. This new approach applying Lyapunov's inequality is, in a worst case, always as good as the standard approach, since $p = 1$ is a feasible solution.

We conclude that the gap of

$$\inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\mathrm{E}\Big[e^{p\theta(A(i,t)-S(i,s))}\Big]\right)^{\frac{1}{p}}\right\} \tag{9.6}$$

and

$$\sum_{i=0}^{s}\Big(\mathrm{E}\Big[e^{\theta(A(i,t)-S(i,s))}\Big]\Big) \tag{9.7}$$

is our possible improvement (in case this gap is strictly > 0). In the following, we call (9.6) the "Lyapunov bound" and the benchmark (9.7) the "Standard Bound".

Remark 9.4. The observation of Remark 9.3 can be restated with the notation as in Definition 2.14 (conventional convolution). We easily see that

$$\begin{aligned}\phi_{A\oslash S\,(s,t)}(\theta) &\overset{(9.3)}{=} \inf_{p\geq 1}\left\{\Big(\phi_{A\oslash S(s,t)}\left(p\theta\right)\Big)^{\frac{1}{p}}\right\}\\ &\overset{\text{(Prop 2.16)}}{\leq} \inf_{p\geq 1}\left\{\left(\left(\phi_A\left(p\theta\right)\circ\phi_S\left(-p\theta\right)\right)(s,t)\right)^{\frac{1}{p}}\right\}\\ &\overset{(*)}{\leq} \left(\phi_A(\theta)\circ\phi_S(-\theta)\right)(s,t),\end{aligned} \tag{9.8}$$

where $(*)$ is the subadditivity of the concave root function, that we have shown in Proposition F.3.

10. Improved Output Bounds for Different Distributions

In the previous Chapter 9, we have introduced a new bound, the so-called "Lyapunov bound" and showed that it is always at least as good as the standard bound. In this chapter, we examine if an improvement is clearly visible or if the lower bound might also be an upper bound.

In all our examples, we assume the service to be of constant rate. We selected the distribution parameters such that

$$\mathrm{P}(A(s,t) > S\,[s,t]) > 0 \qquad \forall 0 \leq s \leq t,$$

(since otherwise, there is no backlog) but also to fulfill

$$\mathrm{E}[A(s,t)] < \mathrm{E}[S(s,t)] \qquad \forall 0 \leq s \leq t,$$

to ensure the stability of the system.

10.1. Exponential Distribution

Let the increments $a(i)$ be independent and exponentially distributed to the parameter λ, i.e., $a(i) \sim Exp\,(\lambda)$, then we obtain for the MGF

$$\mathrm{E}\Big[e^{\theta A(s,t)}\Big] = \mathrm{E}\Big[e^{\theta \sum_{i=s+1}^{t} a(i)}\Big] \overset{\text{(i.i.d.)}}{=} \Big(\mathrm{E}\Big[e^{\theta a(i)}\Big]\Big)^{t-s} = \begin{cases} \left(\frac{\lambda}{\lambda-\theta}\right)^{t-s}, & \text{if } \theta < \lambda \\ \infty, & \text{else.} \end{cases} \tag{10.1}$$

Hence we have for $a(i) \sim \exp(\lambda)$ i.i.d. with $p\theta < \lambda$ and $A(s,t) = \sum_{i=s+1}^{t} a(i)$

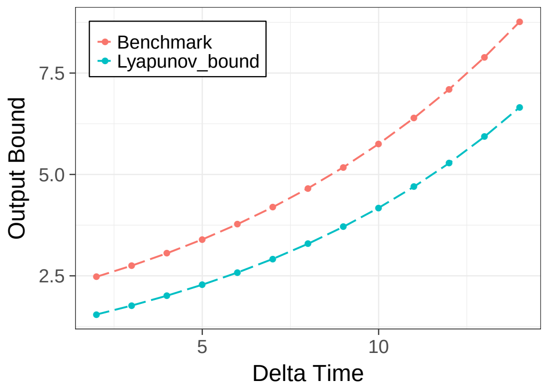
$$\begin{aligned}
\inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\mathrm{E}\Big[e^{p\theta(A(i,t)-S(i,s))}\Big]\right)^{\frac{1}{p}}\right\} &= \inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\mathrm{E}\Big[e^{p\theta(A(i,t))}\Big]\,e^{-p\theta r(s-i)}\right)^{\frac{1}{p}}\right\} \\
&\overset{(10.1)}{=} \inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\left(\frac{\lambda}{\lambda-p\theta}\right)^{t-i} e^{-p\theta r(s-i)}\right)^{\frac{1}{p}}\right\} \\
&= \inf_{p\geq 1}\left\{\left(\left(\frac{\lambda}{\lambda-p\theta}\right)^{t-s}\sum_{i=0}^{s}\left(\frac{\lambda}{\lambda-p\theta}\right)^{s-i} e^{-p\theta r(s-i)}\right)^{\frac{1}{p}}\right\} \\
&= \inf_{p\geq 1}\left\{\left(\left(\frac{\lambda}{\lambda-p\theta}\right)^{t-s}\sum_{i=0}^{s}\left(\frac{\lambda}{\lambda-p\theta}e^{-p\theta r}\right)^{i}\right)^{\frac{1}{p}}\right\}
\end{aligned}$$

$$\stackrel{\text{(geom. series)}}{\leq} \inf_{p\geq 1} \left\{ \left(\left(\frac{\lambda}{\lambda - p\theta}\right)^{t-s} \frac{1-\left(\frac{\lambda}{\lambda-p\theta}e^{-p\theta r}\right)^{s+1}}{1-\frac{\lambda}{\lambda-p\theta}e^{-p\theta r}} \right)^{\frac{1}{p}} \right\}.$$

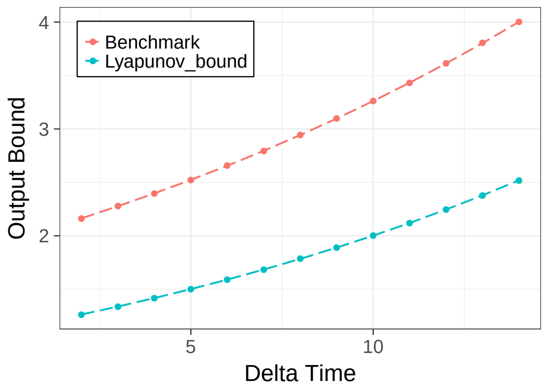
We minimize this bound numerically and compare it to the standard bound

$$\sum_{i=0}^{s} \mathrm{E}\left[e^{\theta(A(i,t)-S(i,s))}\right] = \left(\frac{\lambda}{\lambda-\theta}\right)^{t-s} \frac{1-\left(\frac{\lambda}{\lambda-\theta}e^{-\theta r}\right)^{s+1}}{1-\frac{\lambda}{\lambda-\theta}e^{-\theta r}}.$$

The results depicted in Figure 10.1 indicate that a significant improvement could be accomplished. Some values show a reduction of the function value by 50%. But by comparing the plots in Figure 10.1a and 10.1b we see that this advantage highly depends on the parameter λ.



(a) Exponential distribution with $\lambda = 1$.



(b) Exponential distribution with $\lambda = 2$.

Figure 10.1.: Output bound comparison for the exponential distribution.

Remark 10.1. The exponential distribution plays a central role in queueing theory's M/M/1, the system's description in the "classical queueing system", see for example [Kle75].

10.2. Uniform Distribution

Let the increments $a(i)$ be independent and uniformly distributed, i.e., $a(i) \sim \mathcal{U}(0,1)$, then we have

$$\mathrm{E}\left[e^{\theta A(s,t)}\right] = \mathrm{E}\left[e^{\theta \sum_{i=s+1}^{t} a(i)}\right] \stackrel{\text{(indep.)}}{=} \prod_{i=s+1}^{t} \mathrm{E}\left[e^{\theta a(i)}\right] \stackrel{\text{(i.i.d.)}}{=} \left(\mathrm{E}\left[e^{\theta a(i)}\right]\right)^{t-s} = \left(\frac{e^{\theta}-1}{\theta}\right)^{t-s}. \tag{10.2}$$

This leads to

$$
\begin{aligned}
\inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\mathrm{E}\left[e^{p\theta(A(i,t)-S(i,s))}\right]\right)^{\frac{1}{p}}\right\} &= \inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\mathrm{E}\left[e^{p\theta A(i,t)}\right]e^{-p\theta r(s-i)}\right)^{\frac{1}{p}}\right\} \\
&\overset{(10.2)}{=} \inf_{p\geq 1}\left\{\left(\sum_{i=0}^{s}\left(\frac{e^{p\theta}-1}{p\theta}\right)^{t-i}e^{-p\theta r(s-i)}\right)^{\frac{1}{p}}\right\} \\
&= \inf_{p\geq 1}\left\{\left(\left(\frac{e^{p\theta}-1}{p\theta}\right)^{t-s}\sum_{i=0}^{s}\left(\frac{e^{p\theta}-1}{p\theta}\right)^{s-i}e^{-p\theta r(s-i)}\right)^{\frac{1}{p}}\right\} \\
&= \inf_{p\geq 1}\left\{\left(\left(\frac{e^{p\theta}-1}{p\theta}\right)^{t-s}\sum_{i=0}^{s}\left(\frac{e^{p\theta}-1}{p\theta}e^{-p\theta r}\right)^{i}\right)^{\frac{1}{p}}\right\} \\
&\leq \inf_{p\geq 1}\left\{\left(\left(\frac{e^{p\theta}-1}{p\theta}\right)^{t-s}\frac{1-\left(\frac{e^{p\theta}-1}{p\theta}e^{-p\theta r}\right)^{s+1}}{1-\frac{e^{p\theta}-1}{p\theta}e^{-p\theta r}}\right)^{\frac{1}{p}}\right\}.
\end{aligned}
$$

Again, we minimize this bound numerically and compare it to the standard bound

$$
\sum_{i=0}^{s}\mathrm{E}\left[e^{\theta(A(i,t)-S(i,s))}\right] = \left(\frac{e^{\theta}-1}{\theta}\right)^{t-s}\frac{1-\left(\frac{e^{\theta}-1}{\theta}e^{-\theta r}\right)^{s+1}}{1-\frac{e^{\theta}-1}{\theta}e^{-\theta r}}.
$$

Compared to the exponential distribution's results in Figure 10.1, we see in Figure 10.2 that the output bounds drift apart from the uniform distribution. Thus we conclude that our new approach performs much better for this distribution since the difference increases over time.

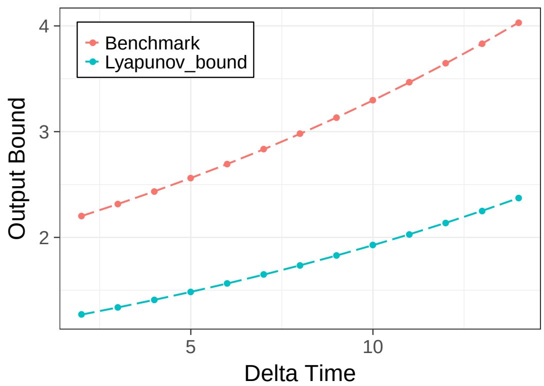


Figure 10.2.: Output bound comparison for the uniform distribution.

10.3. Fractional Brownian Motion

The form of the MGF has been presented in, e.g., [Kel96]. In order to restate the results, we have to define the effective bandwidth at first:

Definition 10.2 (Effective Bandwidth)**.** For the arrival process $A(0,t)$ the *effective bandwidth* is defined as

$$\alpha(\theta,t) := \frac{1}{\theta t}\log\left(\mathrm{E}\Big[e^{\theta A(0,t)}\Big]\right). \tag{10.3}$$

According to [Kel96], if $F(t)$ is a fractional Brownian motion (see Appendix I) with drift $\lambda \geq 0$, variance at $t=1$ equal to σ^2 $\left(\mathrm{var}(F(t)) = \sigma^2 t^{2H}\right)$, then it holds for $A(0,t) = \lambda t + F(t)$ that

$$\alpha(\theta,t) = \lambda + \frac{\sigma^2\theta}{2}t^{2H-1} \tag{10.4}$$

where $H \in (0,1)$ is the *Hurst parameter*. Thus, we obtain

$$\mathrm{E}\Big[e^{\theta A(0,t)}\Big] \overset{(10.3)}{=} e^{\theta t\alpha(\theta,t)} \overset{(10.4)}{=} e^{\lambda\theta t + \frac{(\sigma\theta)^2}{2}t^{2H}}. \tag{10.5}$$

Applying this knowledge to (9.1) and (9.5) is not straightforward since this time, we are dealing with a continuous model. Therefore, we introduce the discretization parameter $\tau > 0$ such that the continuous interval is split into points: we define the points

$$j := \left\lfloor \frac{t-i}{\tau} \right\rfloor \tag{10.6}$$

for all $0 \leq i \leq t$. (For an application of this idea, see for example [CPS13]). Then,

$$\begin{aligned} & j+1 && \overset{(10.6)}{\geq} \frac{t-i}{\tau} \\ \Leftrightarrow\ & (j+1)\tau && \geq t-i \\ \Leftrightarrow\ & i && \geq t-(j+1)\tau \end{aligned}$$

gives us

$$A(i,t) \leq A\left(t-(j+1)\tau, t\right) = A\left(0,(j+1)\tau\right),$$

since the interval becomes larger. The last equality follows from the stationarity of the increments (Theorem I.3). The service has to be lower bounded, so we compute in a similar fashion

$$\begin{aligned} & j && \overset{(10.6)}{\leq} \frac{t-i}{\tau} \\ \Leftrightarrow\ & j\tau && \leq t-i \\ \Leftrightarrow\ & i && \leq t-j\tau. \end{aligned}$$

Then we obtain

$$S(i,s) \geq S\left(t-j\tau, s\right) = S\left(t, s+j\tau\right).$$

We continue with

$$
\begin{aligned}
\mathrm{E}\Big[e^{\theta A'(s,t)}\Big] &\overset{(2.11)}{\leq} \mathrm{E}\Big[e^{\theta A \oslash S\,(s,t)}\Big] \\
&\overset{(2.7)}{=} \mathrm{E}\Big[e^{\theta \max_{0\leq i\leq s}\{A(i,t)-S(i,s)\}}\Big] \\
&\leq \mathrm{E}\left[\max_{\lfloor \frac{t-s}{\tau}\rfloor \leq j \leq \lfloor \frac{t}{\tau}\rfloor} e^{\theta(A(0,(j+1)\tau)-S(t,s+j\tau))}\right] \\
&= \inf_{p\geq 1}\left\{\left(\mathrm{E}\left[\max_{\lfloor \frac{t-s}{\tau}\rfloor \leq j \leq \lfloor \frac{t}{\tau}\rfloor} e^{p\theta(A(0,(j+1)\tau)-S(t,s+j\tau))}\right]\right)^{\frac{1}{p}}\right\} \\
&\leq \inf_{p\geq 1}\left\{\left(\sum_{j=\lfloor \frac{t-s}{\tau}\rfloor}^{\lfloor \frac{t}{\tau}\rfloor} \mathrm{E}\Big[e^{p\theta(A(0,(j+1)\tau)-S(t-j\tau,s))}\Big]\right)^{\frac{1}{p}}\right\} \\
&\overset{(10.5)}{\leq} \inf_{p\geq 1}\left\{\left(\sum_{j=\lfloor \frac{t-s}{\tau}\rfloor}^{\lfloor \frac{t}{\tau}\rfloor} e^{\lambda p\theta((j+1)\tau)+\frac{(\sigma p\theta)^2}{2}((j+1)\tau)^{2H}} e^{-p\theta r(s+j\tau-t)}\right)^{\frac{1}{p}}\right\} \\
&= \inf_{p\geq 1}\left\{\left(e^{p\theta(\lambda\tau+r(t-s))}\sum_{j=\lfloor \frac{t-s}{\tau}\rfloor}^{\lfloor \frac{t}{\tau}\rfloor} e^{\frac{(\sigma p\theta)^2}{2}((j+1)\tau)^{2H}} e^{p\theta(j\tau\lambda-rj\tau)}\right)^{\frac{1}{p}}\right\},
\end{aligned}
$$

with j depending on i and compare it to the standard bound

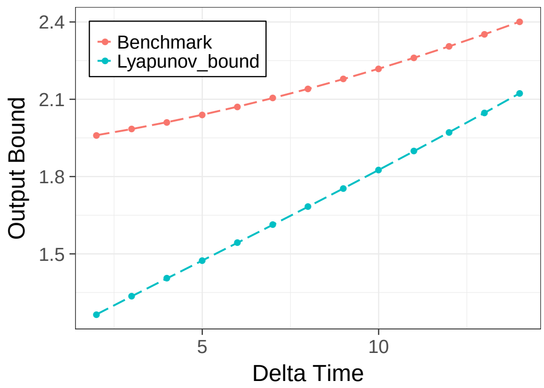
$$
\sum_{j=\lfloor \frac{t-s}{\tau}\rfloor}^{\lfloor \frac{t}{\tau}\rfloor} \mathrm{E}\Big[e^{\theta(A(0,(j+1)\tau)-S(t-j\tau,s))}\Big] \leq e^{\theta(\lambda\tau+r(t-s))}\sum_{j=\lfloor \frac{t-s}{\tau}\rfloor}^{\lfloor \frac{t}{\tau}\rfloor} e^{\frac{(\sigma\theta)^2}{2}((j+1)\tau)^{2H}} e^{\theta(j\tau\lambda-rj\tau)}.
$$

We implemented these bounds; Figure 10.3 depicts the results. The discretization parameter τ is set to 1. As before, we conclude that the Lyapunov bound is able to improve the output bound significantly. But we also observe that for $\lambda = 0$ (Figure 10.3a) or for $H = 0.9$ (Figure 10.3c), the improve decreases quite fast over time whereas for $\lambda = 1$ and $H = 0.7$ (Figure 10.3b), the improvement remains rather stable.

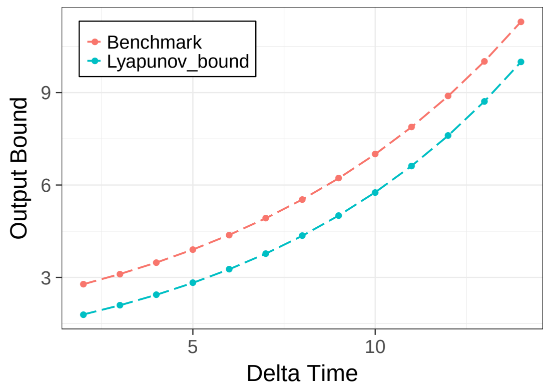
Remark 10.3. The World Wide Web as well as Ethernet traffic have been shown to be of self-similar shape (see [CB97] and [LTWW94]). As we conclude from Theorem I.3 in Appendix I, Fractional Brownian motion shares this property which is the reason why it is a very popular model to simulate this sort of traffic.

10.4. Continuous Markov-Modulated On-Off (MMOO) Model

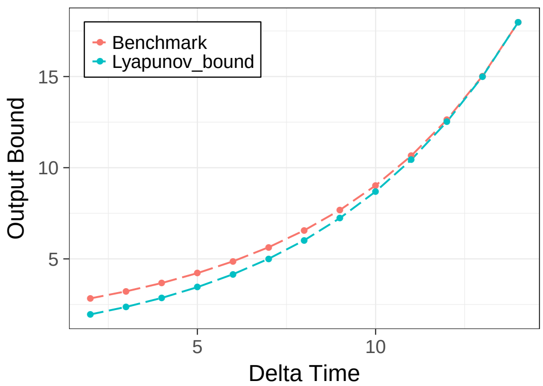
This section follows closely the papers of [CPS13] and [CW96]. Let X_t be a random process of the form $X_t = Z_t \cdot B$, the so-called fluid rate at t, where $B > 0$ is a burst



(a) Fractional Brownian motion with $\lambda = 0$, $\sigma = 1$ and $H = 0.7$.



(b) Fractional Brownian motion with $\lambda = 1$, $\sigma = 1$ and $H = 0.7$.



(c) Fractional Brownian motion with $\lambda = 1$, $\sigma = 1$ and $H = 0.9$.

Figure 10.3.: Output bound comparison for fractional Brownian motion.

parameter and Z_t is a Markov chain with the following transition probabilities $p_{i,j}$, $i, j \in \{0, 1\}$ for the states 0 and 1:

	0	1
0	$1-\mu$	μ
1	λ	$1-\lambda$

with $\mu, \lambda \in (0, 1)$. In other words, the transition matrix is

$$P := \begin{pmatrix} 1-\mu & \mu \\ \lambda & 1-\lambda \end{pmatrix}.$$

Thus, the Markov process is homogeneous, irreducible and aperiodic with positive recurrent states and has, according to Theorem G.7 in the appendix, a unique stationary distribution π that satisfies

$$\lim_{t\to\infty} p_{i,j}^{(t)} = \lim_{t\to\infty} \mathrm{P}(Z_t = j \mid Z_0 = i) = \pi(j) \quad \forall i,j \in \{0,1\}.$$

This can be computed as follows

$$\begin{aligned}
& \pi \cdot P && \overset{!}{=} \pi \\
\Leftrightarrow \quad & \begin{pmatrix}\pi(0) & \pi(1)\end{pmatrix}\begin{pmatrix}1-\mu & \mu \\ \lambda & 1-\lambda\end{pmatrix} && = \begin{pmatrix}\pi(0) & \pi(1)\end{pmatrix} \\
\Leftrightarrow \quad & \begin{pmatrix}\pi(0)(1-\mu)+\pi(1)\lambda & \pi(0)\mu+\pi(1)(1-\lambda)\end{pmatrix} && = \begin{pmatrix}\pi(0) & \pi(1)\end{pmatrix}
\end{aligned}$$

s.t.

$$\pi(0)+\pi(1) = 1$$

which leads to the following linear equations

$$\begin{aligned}
\mathrm{I} \quad & \pi(0)\mu - \pi(1)\lambda && = 0 \\
\mathrm{II} \quad & \pi(0)+\pi(1) && = 1.
\end{aligned}$$

We insert $\pi(0) = 1-\pi(1)$ from II into I and solve for $\pi(1)$ getting

$$\begin{aligned}
& (1-\pi(1))\,\mu - \pi(1)\lambda && = 0 \\
\Leftrightarrow \quad & \pi(1)(\lambda+\mu) && = \mu \\
\Leftrightarrow \quad & \pi(1) && = \frac{\mu}{\lambda+\mu}
\end{aligned}$$

and hence

$$\pi(0) \overset{(\mathrm{II})}{=} 1-\pi(1) = 1 - \frac{\mu}{\lambda+\mu} = \frac{\lambda}{\lambda+\mu}.$$

For further information about Markov chains, check Appendix G. We continue as in [CW96] for the computation of the MGF. The arrivals are defined as

$$A(t) = \int_0^t X_s \mathrm{d}s \tag{10.7}$$

(since we assume a continuous model) and thus the MGF in state i at time t equals

$$\phi_{i,t}(\theta) = \mathrm{E}_i\left[e^{\theta\int_0^t X_s \mathrm{d}s}\right] \tag{10.8}$$

such that

$$\mathrm{E}\left[e^{\theta\int_0^t X_t \mathrm{d}t}\right] = \pi(0)\phi_{0,t}(\theta) + \pi(1)\phi_{1,t}(\theta).$$

This leads to

$$\begin{aligned}
\mathrm{I} \quad \phi_{0,t+\delta}(\theta) &= (1-\mu\delta)\,\phi_{0,t}(\theta) + \mu\delta\phi_{1,t}(\theta) + o(\delta) && \text{(switches to "on" with prob. } \mu) \\
\mathrm{II} \quad \phi_{1,t+\delta}(\theta) &= e^{\theta B\delta}\left(\lambda\delta\phi_{0,t}(\theta) + (1-\lambda\delta)\,\phi_{1,t}(\theta)\right) + o(\delta) && \text{(switches to "off" with prob. } \lambda)
\end{aligned}$$

and by

$$
\begin{aligned}
\lim_{\delta\to 0} \frac{\partial \phi_{0,t+\delta}(\theta)}{\partial \delta} &= \lim_{\delta\to 0} -\mu\phi_{0,t}(\theta) + \mu\phi_{1,t}(\theta) = -\mu\phi_{0,t}(\theta) + \mu\phi_{1,t}(\theta),\\
\lim_{\delta\to 0} \frac{\partial \phi_{1,t+\delta}(\theta)}{\partial \delta} &= \lim_{\delta\to 0} \theta B e^{\theta B\delta} \left(\lambda\delta\phi_{0,t}(\theta) + (1-\lambda\delta)\,\phi_{1,t}(\theta)\right) + e^{\theta B\delta}\left(\lambda\phi_{0,t}(\theta) - \lambda\phi_{1,t}(\theta)\right)\\
&= \theta B\phi_{1,t}(\theta) + \left(\lambda\phi_{0,t}(\theta) - \lambda\phi_{1,t}(\theta)\right) = \lambda\phi_{0,t}(\theta) + (\theta B - \lambda)\,\phi_{1,t}(\theta)
\end{aligned}
$$

we obtain a linear system of differential equations

$$
\begin{pmatrix} \dot{\phi_{0,t}}(\theta) \\ \dot{\phi_{1,t}}(\theta) \end{pmatrix} = \begin{pmatrix} -\mu & \mu \\ \lambda & \theta B - \lambda \end{pmatrix} \begin{pmatrix} \phi_{0,t}(\theta) \\ \phi_{1,t}(\theta) \end{pmatrix}. \tag{10.9}
$$

In order to diagonalize (see therefore Appendix H), we compute the characteristic polynomial of

$$
A := \begin{pmatrix} -\mu & \mu \\ \lambda & \theta B - \lambda \end{pmatrix}
$$

by

$$
\begin{aligned}
\det \begin{pmatrix} -\mu-\omega & \mu \\ \lambda & \theta B - \lambda - \omega \end{pmatrix} & \stackrel{!}{=} 0\\
\Leftrightarrow \quad \omega^2 + (\mu+\lambda-\theta B)\,\omega + (-\mu)\,(\theta B - \lambda) - \mu\lambda &= 0\\
\Leftrightarrow \quad \omega^2 + (\mu+\lambda-\theta B)\,\omega - \mu\theta B &= 0
\end{aligned}
$$

and the roots (the eigenvalues) are

$$
\begin{aligned}
\omega_1 &= -\frac{\mu+\lambda-\theta B}{2} - \frac{1}{2}\sqrt{(\mu+\lambda-\theta B)^2 + 4\mu\theta B},\\
\omega_2 &= -\frac{\mu+\lambda-\theta B}{2} + \frac{1}{2}\sqrt{(\mu+\lambda-\theta B)^2 + 4\mu\theta B}.
\end{aligned} \tag{10.10}
$$

Note that $\omega_1 + \omega_2 = \theta B - \mu - \lambda$ holds.

We determine the first eigenvector and obtain

$$
\begin{aligned}
(A - \omega_1 \mathrm{I})\, v_1 & \stackrel{!}{=} 0\\
\Leftrightarrow \quad \begin{pmatrix} -\mu-\omega_1 & \mu \\ \lambda & \theta B - \lambda - \omega_1 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

leading to

$$
v_1 = \begin{pmatrix} \mu \\ \mu + \omega_1 \end{pmatrix} \qquad\qquad v_2 = \begin{pmatrix} \theta B - \lambda - \omega_1 \\ -\lambda \end{pmatrix}.
$$

After term manipulation and computing the inverse etc., the authors of [CW96] receive

$$
\begin{aligned}
\phi_{0,t}(\theta) &= \frac{\omega_2}{\omega_2 - \omega_1} e^{\omega_1 t} - \frac{\omega_1}{\omega_2 - \omega_1} e^{\omega_2 t},\\
\phi_{1,t}(\theta) &= \frac{\omega_2 - \theta B}{\omega_2 - \omega_1} e^{\omega_1 t} - \frac{\theta B - \omega_1}{\omega_2 - \omega_1} e^{\omega_2 t}.
\end{aligned}
$$

We continue with (10.8) and see that

$$
\begin{aligned}
\mathrm{E}\left[e^{\theta \int_0^t X_t \mathrm{d}t}\right] &= \pi(0)\phi_{0,t}(\theta) + \pi(1)\phi_{1,t}(\theta) \\
&= \frac{\lambda}{\lambda+\mu}\left(\frac{\omega_2}{\omega_2-\omega_1}e^{\omega_1 t} - \frac{\omega_1}{\omega_2-\omega_1}e^{\omega_2 t}\right) \\
&+ \frac{\mu}{\lambda+\mu}\left(\frac{\omega_2-\theta B}{\omega_2-\omega_1}e^{\omega_1 t} - \frac{\theta B-\omega_1}{\omega_2-\omega_1}e^{\omega_2 t}\right) \\
&= \frac{\lambda\omega_2 + \mu\left(\omega_2-\theta B\right)}{\left(\lambda+\mu\right)\left(\omega_2-\omega_1\right)}e^{\omega_1 t} + \frac{-\lambda\omega_1 + \mu\left(\theta B-\omega_1\right)}{\left(\lambda+\mu\right)\left(\omega_2-\omega_1\right)}e^{\omega_2 t}.
\end{aligned}
$$

Note that

$$
\frac{\lambda\omega_2 + \mu\left(\omega_2-\theta B\right)}{\left(\lambda+\mu\right)\left(\omega_2-\omega_1\right)} + \frac{-\lambda\omega_1 + \mu\left(\theta B-\omega_1\right)}{\left(\lambda+\mu\right)\left(\omega_2-\omega_1\right)} = \frac{\lambda\left(\omega_2-\omega_1\right) + \mu\left(\omega_2-\omega_1\right)}{\left(\lambda+\mu\right)\left(\omega_2-\omega_1\right)} = 1,
$$

and hence, by making use of $\omega_1 \leq \omega_2$, it holds true that

$$
\mathrm{E}\left[e^{\theta A(t)}\right] = \mathrm{E}\left[e^{\theta \int_0^t X_t \mathrm{d}t}\right] \leq e^{\omega_2 t} \tag{10.11}
$$

with ω_2 depending on θ.

As in the previous section, we need to discretize this continuous model. Again, we define the points

$$
j := \left\lfloor \frac{t-i}{\tau} \right\rfloor \tag{10.12}
$$

for all $0 \leq i \leq t$. Then it holds that

$$
A(i,t) \leq A\left(t-(j+1)\tau, t\right) = A\left(0,(j+1)\tau\right)
$$

as well as

$$
S(i,s) \geq S\left(t-j\tau, s\right) = S(t, s+j\tau).
$$

Again, the last equality follows by the stationarity (this time, see [CPS13]).

This results in

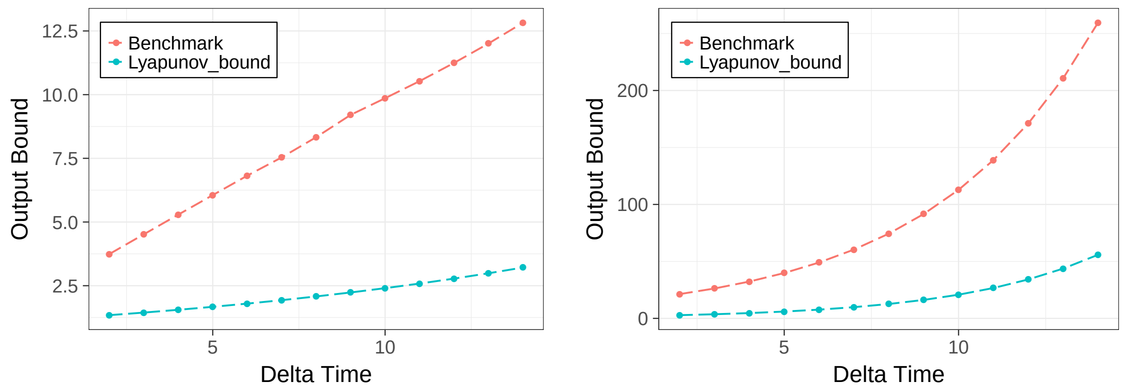
$$
\begin{aligned}
\mathrm{E}\left[e^{\theta A'(s,t)}\right] &\overset{(2.11)}{\leq} \mathrm{E}\left[e^{\theta A \oslash S\,(s,t)}\right] \\
&\overset{(2.7)}{=} \mathrm{E}\left[e^{\theta \max_{0\leq i\leq s}\{A(i,t)-S(i,s)\}}\right] \\
&\leq \mathrm{E}\left[\max_{\left\lfloor\frac{t-s}{\tau}\right\rfloor \leq j \leq \left\lfloor\frac{t}{\tau}\right\rfloor} e^{\theta(A(0,(j+1)\tau)-S(t,s+j\tau))}\right] \\
&= \inf_{p\geq 1}\left\{\left(\mathrm{E}\left[\max_{\left\lfloor\frac{t-s}{\tau}\right\rfloor \leq j \leq \left\lfloor\frac{t}{\tau}\right\rfloor} e^{p\theta(A(0,(j+1)\tau)-S(t,s+j\tau))}\right]\right)^{\frac{1}{p}}\right\} \\
&\leq \inf_{p\geq 1}\left\{\left(\sum_{j=\left\lfloor\frac{t-s}{\tau}\right\rfloor}^{\left\lfloor\frac{t}{\tau}\right\rfloor} \mathrm{E}\left[e^{p\theta(A(0,(j+1)\tau)-S(t,s+j\tau))}\right]\right)^{\frac{1}{p}}\right\}
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(10.11)}{\leq} \inf_{p\geq 1}\left\{\left(\sum_{j=\lfloor\frac{t-s}{\tau}\rfloor}^{\lfloor\frac{t}{\tau}\rfloor} e^{\omega_2(p\theta)(j+1)\tau} e^{-p\theta r(s+j\tau-t)}\right)^{\frac{1}{p}}\right\}\\
&= \inf_{p\geq 1}\left\{\left(e^{\omega_2(p\theta)\tau+p\theta r(t-s)} \sum_{j=\lfloor\frac{t-s}{\tau}\rfloor}^{\lfloor\frac{t}{\tau}\rfloor} e^{\omega_2(p\theta)j\tau-p\theta rj\tau}\right)^{\frac{1}{p}}\right\}\\
&= \inf_{p\geq 1}\left\{\left(e^{\omega_2(p\theta)\tau+p\theta r(t-s)} \sum_{j=\lfloor\frac{t-s}{\tau}\rfloor}^{\lfloor\frac{t}{\tau}\rfloor} \left(e^{\omega_2(p\theta)\tau-p\theta r\tau}\right)^{j}\right)^{\frac{1}{p}}\right\}\\
&= \inf_{p\geq 1}\left\{\left(e^{\omega_2(p\theta)\tau+p\theta r(t-s)} \frac{\left(e^{\omega_2(p\theta)\tau-p\theta r\tau}\right)^{\lfloor\frac{t-s}{\tau}\rfloor} - \left(e^{\omega_2(p\theta)\tau-p\theta r\tau}\right)^{\lfloor\frac{t}{\tau}\rfloor+1}}{1-e^{\omega_2(p\theta)\tau-p\theta r\tau}}\right)^{\frac{1}{p}}\right\}
\end{aligned}
$$

that we compare to the standard bound

$$
\sum_{j=\lfloor\frac{t-s}{\tau}\rfloor}^{\lfloor\frac{t}{\tau}\rfloor} \mathrm{E}\left[e^{\theta(A(0,(j+1)\tau)-S(t,s+j\tau))}\right] \leq e^{\omega_2(\theta)\tau+\theta r(t-s)} \sum_{j=\lfloor\frac{t-s}{\tau}\rfloor}^{\lfloor\frac{t}{\tau}\rfloor} e^{\omega_2(\theta)j\tau+\theta rj\tau}.
$$

Again, we implemented these bounds and visualized the results in Figure 10.4. In each plot, the discretization parameter τ is set to 1. Our conclusion is again very similar to previous plots: we see a strong improvement (again up to approx. 50%) that highly depends on the input parameter that can shrink the gap almost arbitrarily.



(a) MMOO with $r = 3$, $\lambda = 0.4$, $\mu = 0.7$ and $B = 1$. (b) MMOO with $r = 3$, $\lambda = 0.4$, $\mu = 0.7$ and $B = 3$.

Figure 10.4.: Output bound comparison for the MMOO model.

Conclusion and Outlook

This thesis treated several fields of network calculus and thus, a conclusion and possible future work is accordingly.

In the introduction to network calculus we gave all the results necessary for this thesis as well as some examples of how to obtain performance bounds. Afterwards, we were able to show how versatile and useful the Conversion Theorem is by applying it to the the tail-bounds from [MB00] and [VL01]. We have also implemented the MGF-bounds and were able to compare them as far as possible despite some important differences in these models. Due to the fact that these models make only very few assumptions, they are well suited for comparison to DNC which should be tackled in future work.

In the following chapter, we applied this newly gained knowledge to sink trees and observed, in particular, a big performance gap between the SFA and the PMOO-SFA for this topology. This lead us to the conclusion that dependency caused by the analysis has to be avoided by any means necessary due to significant deteriorations of the bounds' qualities. Further research could show how this avoidance could be achieved optimally. We have also shown how much impact parameter optimization can have. Even though we only tried four different approaches, the improvement was very considerable. This field is also open to further examination, for example with other optimization algorithms or by finding criteria for convexity.

In the last part, we have introduced a, to our best knowledge, new modification of the MGF standard output bound and have given several examples where the improvement was significant. It is yet to be shown if this improvement is crucial for real-world applications and how it performs in comparison to other methods that claim to improve the standard approaches.

Bibliography

[BBLO00] Robert R Boorstyn, Almut Burchard, Jörg Liebeherr, and Chaiwat Oottamakorn. Statistical service assurances for traffic scheduling algorithms. *IEEE Journal on Selected Areas in Communications*, 18(12):2651–2664, 2000.

[BCOQ92] François Baccelli, Guy Cohen, Geert Jan Olsder, and Jean-Pierre Quadrat. *Synchronization and linearity: an algebra for discrete event systems.* John Wiley & Sons Ltd, 1992.

[Bec16] Michael A Beck. *Advances in Theory and Applicability of Stochastic Network Calculus.* PhD thesis, TU Kaiserslautern, 2016.

[BJT10] Anne Bouillard, Laurent Jouhet, and Éric Thierry. Tight performance bounds in the worst-case analysis of feed-forward networks. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[BNS16] Steffen Bondorf, Paul Nikolaus, and Jens B Schmitt. Delay Bounds in Feed-Forward Networks - A Fast and Accurate Network Calculus Solution. *CoRR*, abs/1603.02094, 2016.

[BS13] Michael A Beck and Jens B Schmitt. The disco stochastic network calculator version 1.0 - when waiting comes to an end. In *Proceedings of the 7th International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS 2013)*, Dec 2013.

[BS15] Steffen Bondorf and Jens B Schmitt. The DiscoDNC v2 – a comprehensive tool for deterministic network calculus. *EAI Endorsed Transactions on Internet of Things*, 15(4), Dec 2015.

[BS16] Steffen Bondorf and Jens B. Schmitt. Improving cross-traffic bounds in feed-forward networks – there is a job for everyone. In *Proceedings of the 18th International GI/ITG Conference on Measurement, Modelling and Evaluation of Dependable Computer and Communication Systems (MMB & DFT 2016)*, April 2016.

[BT08] Anne Bouillard and Éric Thierry. An algorithmic toolbox for network calculus. *Discrete Event Dynamic Systems*, 18(1):3–49, 2008.

[CB97] Mark E Crovella and Azer Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking (ToN)*, 5(6):835–846, 1997.

[CBL05] Florin Ciucu, Almut Burchard, and Jörg Liebeherr. A network service curve approach for the stochastic analysis of networks. In *Proc. ACM SIGMETRICS 2005*, volume 33, pages 279–290, 2005.

[CCS01] Cheng-Shang Chang, Yuh-ming Chiu, and Wheyming Tina Song. On the performance of multiplexing independent regulated inputs. In *Proc. ACM SIGMETRICS 2001*, pages 184–193, June 2001.

[Cha92a] Cheng-Shang Chang. Stability, queue length and delay. ii. stochastic queueing networks. In *Decision and Control, 1992., Proceedings of the 31st IEEE Conference on*, pages 1005–1010. IEEE, 1992.

[Cha92b] Cheng-Shang Chang. Stability, queue length, and delay of deterministic and stochastic queueing networks. *IEEE Transactions on Automatic Control*, 39(5):913–931, 1992.

[Cha00] Cheng-Shang Chang. *Performance guarantees in communication networks.* Springer-Verlag, London, 2000.

[Che01] Wing-Sum Cheung. Generalizations of Hölder's inequality. *Int. J. Math. Math. Sci*, 26(1):7–10, 2001.

[CLB00] Anna Charny and Jean-Yves Le Boudec. Delay bounds in a network with aggregate scheduling. In *International Workshop on Quality of Future Internet Services*, pages 1–13. Springer, 2000.

[CPS13] Florin Ciucu, Felix Poloczek, and Jens B Schmitt. Sharp bounds in stochastic network calculus. In *Proc. ACM SIGMETRICS 2013*, volume 41, pages 367–368. ACM, Jun. 2013.

[Cru91a] Rene L Cruz. A calculus for network delay, part I: Network elements in isolation. *IEEE Transactions on information theory*, 37(1):114–131, 1991.

[Cru91b] Rene L Cruz. A calculus for network delay, part II: Network analysis. *IEEE Transactions on information theory*, 37(1):132–141, 1991.

[Cru95] Rene L Cruz. Quality of service guarantees in virtual circuit switched networks. *IEEE Journal on Selected areas in Communications*, 13(6):1048–1056, 1995.

[CW96] Costas Courcoubetis and Richard Weber. Buffer overflow asymptotics for a buffer handling many traffic sources. *Journal of Applied Probability*, 33:886–903, 1996.

[Erl09] Agner Krarup Erlang. The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B*, 20(33-39):16, 1909.

[Erl17] Agner Krarup Erlang. Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektrotkeknikeren*, 13:5–13, 1917.

[Fid06] Markus Fidler. An end-to-end probabilistic network calculus with moment generating functions. In *Proc. IEEE IWQoS*, pages 261–270, Jun. 2006.

[Fid10] Markus Fidler. Survey of deterministic and stochastic service curve models in the network calculus. *IEEE Communications Surveys & Tutorials*, 12(1):59–86, 2010.

[FR15] Markus Fidler and Amr Rizk. A guide to the stochastic network calculus. *IEEE Communications Surveys & Tutorials*, 17(1):92–105, 2015.

[FS04] Markus Fidler and Volker Sander. A parameter based admission control for differentiated services networks. *Computer Networks*, 44(4):463–479, 2004.

[GZMS08] Nicos Gollan, Frank A Zdarsky, Ivan Martinovic, and Jens Schmitt. The disco network calculator. In *14th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems (MMB 2008)*, Dortmund, Germany, March 2008. GI/ITG.

[Heu13] Harro Heuser. *Gewöhnliche Differentialgleichungen: Einführung in Lehre und Gebrauch*. Springer-Verlag, 2013.

[HJ61] Robert Hooke and Terry A Jeeves. "Direct Search" Solution of Numerical and Statistical Problems. *Journal of the ACM (JACM)*, 8(2):212–229, 1961.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58:13–30, 1963.

[JL08] Yuming Jiang and Yong Liu. *Stochastic network calculus*, volume 1. Springer, 2008.

[JP02] Jean Jacod and Philip E Protter. *Probability essentials*. Springer, 2002.

[KAT06] Anis Koubaa, Mario Alves, and Eduardo Tovar. Modeling and worst-case dimensioning of cluster-tree wireless sensor networks. In *2006 27th IEEE International Real-Time Systems Symposium (RTSS'06)*, pages 412–421. IEEE, 2006.

[Kel96] Frank Patrick Kelly. Notes on Effective Bandwidths. In Frank Patrick Kelly, Stan Zachary, and Ilze Ziedins, editors, *Stochastic Networks: Theory and Applications*, volume 4 of *Royal Statistical Society Lecture Notes Series*, pages 141–168. Oxford University Press: Oxford, 1996.

[KGV83] Scott Kirkpatrick, C Daniel Gelatt, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[KKK10] Ralf Korn, Elke Korn, and Gerald Kroisandt. *Monte Carlo methods and models in finance and insurance.* Chapman & Hall / CRC Financial Mathematics Series, London, 2010.

[Kle75] Leonard Kleinrock. Queueing systems, volume I: theory. 1975.

[Kur92] Jim Kurose. *On computing per-session performance bounds in high-speed multi-hop computer networks*, volume 20. ACM, 1992.

[LBL07] Chengzhi Li, Almut Burchard, and Jörg Liebeherr. A network calculus with effective bandwidth. *IEEE/ACM Transactions on Networking (ToN)*, 15(6):1442–1453, 2007.

[LBT96] Jean-Yves Le Boudec and P Thiran. Network calculus made easy. *preprint*, 1996.

[LFV07] Jörg Liebeherr, Markus Fidler, and Shahrokh Valaee. A min-plus system interpretation of bandwidth estimation. In *IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications*, pages 1127–1135. IEEE, 2007.

[LMMS06] Luciano Lenzini, Linda Martorini, Enzo Mingozzi, and Giovanni Stea. Tight end-to-end per-flow delay bounds in FIFO multiplexing sink-tree networks. *Performance Evaluation*, 63(9):956–987, 2006.

[LT01] Jean-Yves Le Boudec and Patrick Thiran. *Network calculus: a theory of deterministic queuing systems for the internet.* Springer-Verlag, New York, 2001.

[LTWW94] Will E Leland, Murad S Taqqu, Walter Willinger, and Daniel V Wilson. On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking (ToN)*, 2(1):1–15, 1994.

[LWF96] Jörg Liebherr, DE Werge, and Domenico Ferrari. Exact admission control for networks with a bounded delay service. *IEEE/ACM Transactions on Networking (ToN)*, 4:885–901, 1996.

[LZ10] Chengzhi Li and Wei Zhao. Stochastic performance analysis of non-feedforward networks. *Telecommunication Systems*, 43(3-4):237–252, 2010.

[MB00] Laurent Massoulié and Anthony Busson. Stochastic majorization of aggregates of leaky bucket-constrained traffic streams, 2000.

[MP06] Shiwen Mao and Shivendra S Panwar. A survey of envelope processes and their applications in quality of service provisioning. *IEEE Communications Surveys & Tutorials*, 8(3):2–20, 2006.

[MVN68] Benoit B Mandelbrot and John W Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.

[NM65] John A Nelder and Roger Mead. A simplex method for function minimization. *Computer journal*, 7:308–313, 1965.

[PG93] Abhay K Parekh and Robert G Gallager. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM Transactions on Networking (ToN)*, 1(3):344–357, 1993.

[PG94] Abhay K Parekh and Robert G Gallager. A generalized processor sharing approach to flow control in integrated services networks: the multiple-node case. *IEEE/ACM Transactions on Networking (ToN)*, 2(2):137–150, 1994.

[Ros06] Sheldon M Ross. *A first course in probability*, volume 7. Pearson Prentice Hall Upper Saddle River, NJ, 2006.

[RSF07] Frédéric Ridouard, Jean-Luc Scharbarg, and Christian Fraboul. Stochastic network calculus for end-to-end delays distribution evaluation on an avionics switched ethernet. In *2007 5th IEEE International Conference on Industrial Informatics*, volume 1, pages 559–564. IEEE, 2007.

[RVC01] Eric Rosen, Arun Viswanathan, and Ross Callon. Multiprotocol label switching architecture. Technical report, 2001.

[SJH06] Tor Skeie, Svein Johannessen, and Oyvind Holmeide. Timeliness of real-time ip communication in switched industrial ethernet networks. *IEEE Transactions on Industrial Informatics*, 2(1):25–39, 2006.

[SR05] Jens B Schmitt and Utz Roedig. Sensor Network Calculus - A Framework for Worst Case Analysis. In *Proceedings of IEEE/ACM International Conference on Distributed Computing in Sensor Systems (DCOSS'05), Marina del Rey, USA*, pages 141–154. Springer, LNCS 3560, June 2005.

[SZ06] Jens B Schmitt and Frank A Zdarsky. The DISCO Network Calculator - A Toolbox for Worst Case Analysis. In *Proceedings of the First International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS'06), Pisa, Italy.* ACM, November 2006.

[SZF08] Jens B Schmitt, Frank A Zdarsky, and Markus Fidler. Delay bounds under arbitrary multiplexing: When network calculus leaves you in the lurch ... In *27th IEEE International Conference on Computer Communications (INFOCOM 2008)*, Phoenix, AZ, USA, April 2008.

[SZM08] Jens Schmitt, Frank A Zdarsky, and Ivan Martinovic. Improving performance bounds in feed-forward networks by paying multiplexing only once. In *14th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems (MMB 2008)*, Dortmund, Germany, March 2008. GI/ITG.

[VL01] Milan Vojnović and Jean-Yves Le Boudec. Bounds for independent regulated inputs multiplexed in a service curve network element. In *Proc. GLOBECOM*, San Antonio, TX, Nov. 2001.

[VLB02] Milan Vojnovic and J-Y Le Boudec. Stochastic analysis of some expedited forwarding networks. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 1004–1013. IEEE, 2002.

[WT06] Ernesto Wandeler and Lothar Thiele. Real-time calculus (RTC) toolbox, 2006.

[YS93] Opher Yaron and Moshe Sidi. Performance and stability of communication networks via robust exponential bounds. *IEEE/ACM Transactions on Networking (ToN)*, 1(3):372–385, 1993.

[ZBHB16] T. Zhu, D. S. Berger, and M. Harchol-Balter. SNC-meister: Admitting more tenants with tail latency SLOs. In *ACM SoCC (Symposium on Cloud Computing)*, 2016.

Appendix

A. Bound Computation for the 2- and n-Chain

A.1. Proposition 2.41

$$
\begin{aligned}
\mathrm{P}(d(t) > T) &\overset{(2.13)}{\leq} \mathrm{E}\left[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o.}}(t+T,t)}\right] \\
&\overset{(2.21)}{\leq} \sum_{k=0}^{t+T} \mathrm{E}\left[e^{\theta(A_{\mathrm{foi}}(k,t) - S_{\mathrm{l.o.}}(k,t+T))}\right] \\
&= \sum_{k=0}^{t+T} \mathrm{E}\left[e^{\theta\left(A_1(k,t) - \left([S_1 - A_2]^+ \otimes [S_2 - (A_2 \oslash S_1)]^+\right)(k,t+T)\right)}\right] \\
&\overset{\text{(indep.)}}{=} \sum_{k=0}^{t+T} \mathrm{E}\left[e^{\theta A_1(k,t)}\right] E\left[e^{-\theta\left([S_1 - A_2]^+ \otimes [S_2 - (A_2 \oslash S_1)]^+\right)(k,t+T)}\right] \\
&\overset{(2.21)}{\leq} \sum_{k=0}^{t+T} \left(E\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{-\theta\left([S_1 - A_2]^+(k,i) + [S_2 - (A_2 \oslash S_1)]^+(i,t+T)\right)}\right] \right) \\
&= \sum_{k=0}^{t+T} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{-\theta[S_1 - A_2]^+(k,i)} e^{-\theta[S_2 - (A_2 \oslash S_1)]^+(i,t+T)}\right] \right).
\end{aligned}
$$

Now, we apply Hölder's inequality:

$$
\begin{aligned}
\ldots &\overset{(2.20)}{\leq} \sum_{k=0}^{t+T} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{-p\theta[S_1 - A_2]^+(k,i)}\right]^{\frac{1}{p}} E\left[e^{-q\theta[S_2 - (A_2 \oslash S_1)]^+(i,t+T)}\right]^{\frac{1}{q}} \right) \\
&= \sum_{k=0}^{t+T} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{p\theta \min\{0, A_2 - S_1\}(k,i)}\right]^{\frac{1}{p}} E\left[e^{q\theta \min\{0, (A_2 \oslash S_1) - S_2\}(i,t+T)}\right]^{\frac{1}{q}} \right) \\
&\leq \sum_{k=0}^{t+T} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{p\theta A_2 - p\theta S_1(k,i)}\right]^{\frac{1}{p}} E\left[e^{q\theta(A_2 \oslash S_1) - q\theta S_2(i,t+T)}\right]^{\frac{1}{q}} \right) \\
&\overset{\text{(indep.)}}{=} \sum_{k=0}^{t+T} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{p\theta A_2(k,i)}\right]^{\frac{1}{p}} E\left[e^{-p\theta S_1(k,i)}\right]^{\frac{1}{p}} \right. \\
&\qquad \left. \cdot E\left[e^{q\theta(A_2 \oslash S_1)(i,t+T)}\right]^{\frac{1}{q}} E\left[e^{-q\theta S_2(i,t+T)}\right]^{\frac{1}{q}} \right) \\
&\overset{(2.21)}{\leq} \sum_{k=0}^{t+T} \left(\mathrm{E}\left[e^{\theta A_1(k,t)}\right] \sum_{i=k}^{t+T} E\left[e^{p\theta A_2(k,i)}\right]^{\frac{1}{p}} E\left[e^{-p\theta S_1(k,i)}\right]^{\frac{1}{p}} \right.
\end{aligned}
$$

$$
\begin{aligned}
&\cdot\left(\sum_{j=0}^{i} E\left[e^{q\theta(A_2(j,t)-S_1(j,i))}\right]^{\frac{1}{q}}\right) E\left[e^{-q\theta S_2(i,t+T)}\right]^{\frac{1}{q}}\Bigg) \\
&\overset{\text{(indep.)}}{=} \sum_{k=0}^{t+T}\Bigg(\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] \sum_{i=k}^{t+T} E\left[e^{p\theta A_2(k,i)}\right]^{\frac{1}{p}} E\left[e^{-p\theta S_1(k,i)}\right]^{\frac{1}{p}} \\
&\cdot\left(\sum_{j=0}^{i} E\left[e^{q\theta A_2(j,t)}\right]^{\frac{1}{q}} E\left[e^{-q\theta S_1(j,i)}\right]^{\frac{1}{q}}\right) E\left[e^{-q\theta S_2(i,t+T)}\right]^{\frac{1}{q}}\Bigg),
\end{aligned}
$$

with

$$\frac{1}{p}+\frac{1}{q}=1.$$

A.2. Proposition 2.42

$$
\begin{aligned}
\mathrm{P}(d(t)>T) &\le \sum_{k=0}^{t+T}\mathrm{E}\Big[e^{\theta(A_{\text{foi}}(k,t)-S_{\text{l.o.}}(k,t+T))}\Big] \\
&= \sum_{k=0}^{t+T} E\left[e^{\theta\left(A_1(k,t)-[S_1\otimes S_2(k,t+T)-A_2(k,t+T)]^+\right)}\right] \\
&= \sum_{k=0}^{t+T}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] E\left[e^{-\theta[S_1\otimes S_2(k,t+T)-A_2(k,t+T)]^+}\right] \\
&\le \sum_{k=0}^{t+T}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] E\left[e^{\theta(A_2(k,t+T)-S_1\otimes S_2(k,t+T))}\right] \\
&= \sum_{k=0}^{t+T}\mathrm{E}\Big[e^{\theta A_1(k,t)}\Big] E\left[e^{\theta(A(k,t+T))}\right] E\left[e^{-\theta S_1\otimes S_2(k,t+T)}\right] \\
&\overset{(2.21)}{\le} \sum_{k=0}^{t+T}\left(E\left[e^{\theta A_1(k,t)}\right] E\left[e^{\theta A_2(k,t+T)}\right]\sum_{i=k}^{t+T} E\left[e^{-\theta(S_1(k,i)+S_2(i,t+T))}\right]\right) \\
&= \sum_{k=0}^{t+T}\left(E\left[e^{\theta A_1(k,t)}\right] E\left[e^{\theta A_2(k,t+T)}\right]\sum_{i=k}^{t+T} E\left[e^{-\theta S_1(k,i)}\right] E\left[e^{-\theta S_2(i,t+T)}\right]\right).
\end{aligned}
$$

A.3. Proposition 2.43

$$
\begin{aligned}
\phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\le} \mathrm{E}\Big[e^{\theta A_{\text{foi}}\oslash S_{\text{l.o.}}\,(s,t)}\Big] \\
&\overset{(2.7)}{=} \mathrm{E}\Big[e^{\theta\max_{0\le k_0\le s}\{A_{\text{foi}}(k_0,t)-S_{\text{l.o.}}(k_0,s)\}}\Big] \\
&\overset{(2.21)}{\le} \sum_{k_0=0}^{s}\mathrm{E}\Big[e^{\theta(A_{\text{foi}}(k_0,t)-S_{\text{l.o.}}(k_0,s))}\Big]
\end{aligned}
$$

$$\overset{\text{(indep.)}}{=} \sum_{k_0=0}^{s} \left(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{-\theta\left(\left[S_1 - \sum_{j=2}^{m} A_j\right]^+ \otimes \left[S_2 - \left(\sum_{j=2}^{m} A_j\right) \oslash S_1\right]^+ \otimes \right.} \right.\right.$$

$$\left.\left.\cdots^{\left.\cdots \otimes \left[S_n - \left(\left(\left(\sum_{j=2}^{m} A_j\right)\oslash S_1\right)\ldots\right)\oslash S_{n-1}\right]^+ (k_0,s)\right)} \right]\right)$$

$$\overset{(2.17)}{\leq} \sum_{k_0=0}^{s} \left(E\left[e^{\theta A_1(k_0,t)}\right] \sum_{k_0 \leq k_1 \leq s} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq s} E\left[e^{-\theta\left(\left[S_1 - \sum_{j=2}^{m} A_j\right]^+ (k_0,k_1) + \right.} \right.\right.$$

$$\left.\left.\cdots^{\left.\cdots + \left[S_n - \left(\left(\left(\sum_{j=2}^{m} A_j\right)\oslash S_1\right)\ldots\right)\oslash S_{n-1}\right]^+ (k_{n-1},s)\right)} \right]\right)$$

All of the factors are stochastically dependent and hence we apply the generalized Hölder inequality:

$$\cdots \overset{(2.22)}{\leq} \sum_{k_0=0}^{s} \left(E\left[e^{\theta A_1(k_0,t)}\right] \sum_{k_0 \leq k_1 \leq s} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq s} E\left[e^{-p_1\theta\left[S_1 - \sum_{j=2}^{m} A_j\right]^+ (k_0,k_1)} \right]^{\frac{1}{p_1}} \right.$$

$$\left. \cdots E\left[e^{p_n\theta\left[S_n - \left(\left(\left(\sum_{j=2}^{m} A_j\right)\oslash S_1\right)\ldots\right)\oslash S_{n-1}\right]^+ (k_{n-1},s)} \right]^{\frac{1}{p_n}} \right)$$

$$\leq \sum_{k_0=0}^{s} \left(E\left[e^{\theta A_1(k_0,t)}\right] \sum_{k_0 \leq k_1 \leq s} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq s} E\left[e^{p_1\theta\left(\sum_{j=2}^{m} A_j - S_1\right)(k_0,k_1)} \right]^{\frac{1}{p_1}} \right.$$

$$\left. \cdots E\left[e^{p_n\theta\left(\left(\left(\left(\left(\sum_{j=2}^{m} A_j\right)\oslash S_1\right)\ldots\right)\oslash S_{n-1}\right) - S_n\right)(k_{n-1},s)} \right]^{\frac{1}{p_n}} \right)$$

$$= \sum_{k_0=0}^{s} \left(E\left[e^{\theta A_1(k_0,t)}\right] \sum_{k_0 \leq k_1 \leq s} \cdots \sum_{k_{n-2} \leq k_{n-1} \leq s} E\left[e^{p_1\theta \sum_{j=2}^{m} A_j(k_0,k_1)} e^{-p_1\theta S_1(k_0,k_1)} \right]^{\frac{1}{p_1}} \right.$$

$$\left. \cdots E\left[e^{p_n\theta\left(\left(\left(\left(\sum_{j=2}^{m} A_j\right)\oslash S_1\right)\ldots\right)\oslash S_{n-1}\right)(k_{n-1},s)} e^{-p_n\theta S_n(k_{n-1},s)} \right]^{\frac{1}{p_n}} \right),$$

with

$$\sum_{i=1}^{n} \frac{1}{p_i} = 1.$$

A.4. Proposition 2.44

$$
\begin{aligned}
\phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\leq} \mathrm{E}\Big[e^{\theta A_{\mathrm{foi}} \oslash S_{\mathrm{l.o.}}\,(s,t)}\Big] \\
&\overset{(2.7)}{=} E\left[e^{\theta \max_{0\leq k_0 \leq s}\{A_{\mathrm{foi}}(k_0,t)-S_{\mathrm{l.o.}}(k_0,s)\}}\right] \\
&\overset{(2.21)}{\leq} \sum_{k_0=0}^{s} E\left[e^{\theta(A_{\mathrm{foi}}(k_0,t)-S_{\mathrm{l.o.}}(k_0,s))}\right] \\
&\overset{(\mathrm{indep.})}{=} \sum_{k_0=0}^{s}\left(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{-\theta\left[\otimes_{i=1}^{n} S_i(k_0,s)-\sum_{j=2}^{s} A_j(k_0,s)\right]^+}\right]\right) \\
&= \sum_{k_0=0}^{s}\left(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{\theta \min\left\{0,\left[\sum_{j=2}^{s}(A_j(k_0,s))-\otimes_{i=1}^{n} S_i(k_0,s)\right]^+\right\}}\right]\right) \\
&\leq \sum_{k_0=0}^{s}\left(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{\theta \sum_{j=2}^{s}(A_j(k_0,s))-\theta\otimes_{i=1}^{n} S_i(k_0,s)}\right]\right) \\
&\overset{(\mathrm{indep.})}{=} \sum_{k_0=0}^{s}\left(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{\theta \sum_{j=2}^{s} A_j(k_0,s)}\right] E\left[e^{-\theta\otimes_{i=1}^{n} S_i(k_0,s)}\right]\right) \\
&\overset{(2.17)}{\leq} \sum_{k_0=0}^{s}\Bigg(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{\theta \sum_{j=2}^{s} A_j(k_0,s)}\right] \\
&\qquad \sum_{k_0\leq k_1\leq s} \cdots \sum_{k_{n-2}\leq k_{n-1}\leq s} E\left[e^{-\theta(S_1(k_0,k_1)+\cdots+S_n(k_{n-1},s))}\right]\Bigg) \\
&\overset{(\mathrm{indep.})}{=} \sum_{k_0=0}^{s}\Bigg(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{\theta \sum_{j=2}^{s} A_j(k_0,s)}\right] \\
&\qquad \sum_{k_0\leq k_1\leq s} \cdots \sum_{k_{n-2}\leq k_{n-1}\leq s} \left(E\left[e^{-\theta S_1(k_0,k_1)}\right] \cdots E\left[e^{-\theta S_n(k_{n-1},s)}\right]\right)\Bigg). \\
&= \sum_{k_0=0}^{s}\Bigg(E\left[e^{\theta A_1(k_0,t)}\right] E\left[e^{\theta \sum_{j=2}^{s} A_j(k_0,s)}\right] \\
&\qquad \sum_{k_0\leq k_1\leq \cdots\leq k_{n-1}\leq s} \left(E\left[e^{-\theta S_1(k_0,k_1)}\right] \cdots E\left[e^{-\theta S_n(k_{n-1},s)}\right]\right)\Bigg).
\end{aligned}
$$

B. Checking the Integral Calculator

In this chapter, we would like to focus on the computation of the integral from Subsection 4.3.2 by the integral-calculator since it provides us with the steps to check the result.

Therefore, we formulate the integral in a more general form and put the result into a claim. The "Proof" is going to contain all the demonstrated steps from the website as well as additional steps in case we consider them to be necessary.

Claim B.1. *For a, s, $t \geq 0$ and $x \in [0,1]$, it holds that*

$$\int e^{t\sqrt{-a\log(x)s}}\mathrm{d}x = xe^{t\sqrt{-as\log(x)}} - \frac{1}{2}t\sqrt{\pi as}e^{\frac{ast^2}{4}}\operatorname{erf}\left(\sqrt{-\log(x)} - \frac{t\sqrt{as}}{2}\right). \tag{B.1}$$

Remark B.2. We would like to mention the general form of the integral again. Since the $\sum_{i=1}^{n}\sigma_i^2$ as well $\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2$ only play the role of a factor of the integration variable, we abstract both of them by the variable s. The same holds for the factor 2, which is replaced by a. Calling θ t and the integration variable x instead of ε should not cause additional confusion either. Keep in mind that it is crucial to bound x to provide a solution in $\mathbb{R}$. It is reasonable since the integration limits in the Conversion Theorem are 0 and 1 anyway.

Proof. In

$$\int e^{t\sqrt{-a\log(x)s}}\mathrm{d}x$$

we substitute $u = \phi^{-1}(x) = \sqrt{-\log(x)} \quad \Rightarrow x = \phi(u) = e^{-u^2}$ and thus by integration by substitution

$$\int f(x)\mathrm{d}x = \int f\left(\phi(u)\right)\cdot\phi'(u)\mathrm{d}u\bigg|_{u=\phi^{-1}(x)}$$

we obtain for $\phi'(u) = -2ue^{-u^2}$

$$\int e^{t\sqrt{-a\log(x)s}}\mathrm{d}x = \int e^{tu\sqrt{as}}\left(-2u\right)e^{-u^2}\mathrm{d}u = -2\int ue^{-u^2+tu\sqrt{as}}\mathrm{d}u. \tag{B.2}$$

We continue with $\int ue^{-u^2+tu\sqrt{as}}\mathrm{d}u$ and follow the integral calculator by algebraic manipulation

$$u = -\frac{1}{2}\left(-2u + t\sqrt{as}\right) + \frac{1}{2}t\sqrt{as}$$

and thus

$$\begin{aligned}\int ue^{-u^2+tu\sqrt{as}}\mathrm{d}u &= \int -\frac{1}{2}\left(-2u+t\sqrt{as}\right)e^{-u^2+tu\sqrt{as}}\mathrm{d}u + \int \frac{1}{2}t\sqrt{as}e^{-u^2+tu\sqrt{as}}\mathrm{d}u \\ &= \frac{1}{2}\int\left(2u-t\sqrt{as}\right)e^{-u^2+tu\sqrt{as}}\mathrm{d}u + \frac{1}{2}t\sqrt{as}\int e^{-u^2+tu\sqrt{as}}\mathrm{d}u.\end{aligned} \tag{B.3}$$

We continue with the first summand of (B.3), $\int\left(2u-t\sqrt{as}\right)e^{-u^2+tu\sqrt{as}}\mathrm{d}u$. Substituting $v=\psi(u)=-u^2+tu\sqrt{as}\Rightarrow\psi'(u)=-2u+t\sqrt{as}$, we apply again the integration by substitution (this time the other way around)

$$\int f\left(\psi(u)\right)\cdot\psi'(u)\mathrm{d}u\bigg|_{v=\phi(u)} = \int f\left(v\right)\mathrm{d}v$$

and get

$$\int\left(2u-t\sqrt{as}\right)e^{-u^2+tu\sqrt{as}}\mathrm{d}u = -\int e^v\mathrm{d}v = -e^v = -e^{-u^2+tu\sqrt{as}}. \tag{B.4}$$

Next, we move on with the second summand of (B.3), $\int e^{-u^2+tu\sqrt{as}}\mathrm{d}u$. We complete the square

$$-u^2+tu\sqrt{as} = -\left(u^2-ut\sqrt{as}\right) = -\left(u-\frac{t\sqrt{as}}{2}\right)^2+\left(\frac{t\sqrt{as}}{2}\right)^2$$

which yields

$$\int e^{-u^2+tu\sqrt{as}}\mathrm{d}u = \int e^{-\left(u-\frac{t\sqrt{as}}{2}\right)^2+\frac{t^2as}{4}}\mathrm{d}u.$$

Once again, we use the integration by substitution

$$\int f\left(\psi(u)\right)\cdot\psi'(u)\mathrm{d}u\bigg|_{v=\phi(u)} = \int f\left(v\right)\mathrm{d}v$$

by substituting $v=\xi(u)=u-\frac{t\sqrt{as}}{2}\Rightarrow\xi'(u)=1$ and hence

$$\int e^{-\left(u-\frac{t\sqrt{as}}{2}\right)^2+\frac{t^2as}{4}}\mathrm{d}u = e^{\frac{t^2as}{4}}\int e^{-v^2}\mathrm{d}v = e^{\frac{t^2as}{4}}\frac{\sqrt{\pi}}{2}\int\frac{2}{\sqrt{\pi}}e^{-v^2}\mathrm{d}v = \frac{\sqrt{\pi}}{2}e^{\frac{t^2as}{4}}\,\mathrm{erf}(v)\,.$$

Since the integral is now solved, we can undo the substitution:

$$\int e^{-\left(u-\frac{t\sqrt{as}}{2}\right)^2+\frac{t^2as}{4}}\mathrm{d}u = \frac{\sqrt{\pi}}{2}e^{\frac{t^2as}{4}}\,\mathrm{erf}\left(u-\frac{t\sqrt{as}}{2}\right). \tag{B.5}$$

Subsequently, we can solve (B.3):

$$\int ue^{-u^2+tu\sqrt{as}}\mathrm{d}u \overset{\text{(B.3)}}{=} \frac{1}{2}\int \left(2u - t\sqrt{as}\right) e^{-u^2+tu\sqrt{as}}\mathrm{d}u + \frac{1}{2}t\sqrt{as}\int e^{-u^2+tu\sqrt{as}}\mathrm{d}u$$
$$\overset{\text{(B.4),(B.5)}}{=} -\frac{1}{2}e^{-u^2+tu\sqrt{as}} + \frac{1}{2}t\sqrt{as}\frac{\sqrt{\pi}}{2}e^{\frac{t^2as}{4}}\operatorname{erf}\left(u - \frac{t\sqrt{as}}{2}\right)$$
$$= -\frac{1}{2}e^{-u^2+tu\sqrt{as}} + \frac{1}{4}t\sqrt{\pi as}e^{\frac{t^2as}{4}}\operatorname{erf}\left(u - \frac{t\sqrt{as}}{2}\right) \tag{B.6}$$

and thus we obtain by (B.2):

$$\int e^{t\sqrt{-a\log(x)s}}\mathrm{d}x \overset{\text{(B.2)}}{=} -2\int ue^{-u^2+tu\sqrt{as}}\mathrm{d}u$$
$$\overset{\text{(B.6)}}{=} e^{-u^2+tu\sqrt{as}} - \frac{1}{2}t\sqrt{\pi as}e^{\frac{t^2as}{4}}\operatorname{erf}\left(u - \frac{t\sqrt{as}}{2}\right) \tag{B.7}$$

At last, we undo the substitution $u = \sqrt{-\log(x)}$:

$$\int e^{t\sqrt{-a\log(x)s}}\mathrm{d}x \overset{\text{(B.7)}}{=} e^{-u^2}e^{tu\sqrt{as}} - \frac{1}{2}t\sqrt{\pi as}e^{\frac{t^2as}{4}}\operatorname{erf}\left(u - \frac{t\sqrt{as}}{2}\right)$$
$$= xe^{t\sqrt{-as\log(x)}} - \frac{1}{2}t\sqrt{\pi as}e^{\frac{t^2as}{4}}\operatorname{erf}\left(\sqrt{-\log(x)} - \frac{t\sqrt{as}}{2}\right)$$

which is exactly what we claimed. $\square$

C. Proof of Theorem 5.5

As we have already mentioned, we repeat the steps in [VL01] and replace $q(0)$ by $q(t)$. Let us give the prerequisites of the proof at first and then go to the single steps.

Prerequisites

Lemma C.1. *For* $q(t) = A(s,t) - A'(s,t)$, $\sum_{i=1}^{n} g_i \leq 1$ *and*

$$q_i(t) \coloneqq \sum_{0<s\leq t} \{A_i(s,t) - g_i\beta(t-s)\}$$

it holds that

$$q(t) \leq \sum_{i=1}^{n} q_i(t). \tag{C.1}$$

Proof. See [VL01]. □

Next, we need Theorem 2 from [Hoe63]:

Theorem C.2. *If* $X_1 \ldots X_n$ *are independent and* $a_i \leq X_i \leq b_i$*, then for* $t > 0$

$$\mathrm{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \geq t\right) \leq e^{-2n^2t^2/\sum_{i=1}^{n}(b_i-a_i)^2}. \tag{C.2}$$

Proof. See [Hoe63]. □

Proof

$$
\begin{aligned}
\mathrm{P}(q(t) > \gamma) &\overset{(\mathrm{C.1})}{\leq} \mathrm{P}\left(\sum_{i=1}^{I} q_i(t) > \gamma\right) \\
&= \mathrm{P}\left(\sum_{i=1}^{n} q_i(t) - \sum_{i=1}^{I} \mathrm{E}[q_i(t)] > \gamma - \sum_{i=1}^{n} \mathrm{E}[q_i(t)]\right) \\
&= \mathrm{P}\left(\frac{1}{n}\sum_{i=1}^{I} q_i(t) - \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[q_i(t)] > \frac{1}{n}\left(\gamma - \sum_{i=1}^{n} \mathrm{E}[q_i(t)]\right)\right) \\
&\overset{(\mathrm{C.2})}{\leq} \exp\left(-\frac{2n^2\left(\frac{1}{n}\left(\gamma - \sum_{i=1}^{n} \mathrm{E}[q_i(t)]\right)\right)^2}{\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right) \\
&= \exp\left(-\frac{2\left(\gamma - \sum_{i=1}^{n} \mathrm{E}[q_i(t)]\right)^2}{\sum_{i=1}^{n} v\left(\alpha_i, g_i\beta\right)^2}\right) \\
&\leq \exp\left(-\frac{2\left(\gamma - \sum_{i=1}^{n} \bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)\right)^2}{\sum_{i=1}^{I} v\left(\alpha_i, g_i\beta\right)^2}\right),
\end{aligned}
$$

where we have utilized in the last step that $\mathrm{E}[q_i(t)]$ is upper bounded by $\bar{\alpha}_i h\left(\alpha_i, g_i\beta\right)$.

D. Sink Tree Bounds

D.1. Output Bound for SFA

$$
\begin{aligned}
\phi_{A'(s,t)}(\theta) &\overset{(2.14)}{\leq} \mathrm{E}\Big[e^{\theta(A_{\mathrm{foi}} \oslash S_{\mathrm{l.o.}}\,(s,t))}\Big] \\
&\overset{(2.7)}{=} \mathrm{E}\Big[e^{\theta \max_{0 \leq k \leq s}\{A_{\mathrm{foi}}(k,t) - S_{\mathrm{l.o.}}(k,s)\}}\Big] \\
&\overset{(2.21)}{\leq} \sum_{k=0}^{s} E\left[e^{\theta(A_{\mathrm{foi}}(k,t) - S_{\mathrm{l.o.}}(k,s))}\right] \\
&= \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right] \\
&\quad \cdot E\Big[e^{-\theta\big([S_1 - A_1]^+ \otimes [S_2 - (A_2 + A_1 \oslash S_1)]^+ \otimes} \\
&\quad \cdot ^{\cdots \otimes \big[S_n - \big(A_n + A_{n-1} \oslash S_{n-1} + \cdots + \big(\big((A_1 \oslash S_1) \oslash [S_2 - A_2]^+\big) \oslash \cdots \oslash [S_{n-1} - (A_2 + \cdots + A_{n-1})]^+\big)\big)\big]^+ (k,s)\big)}\Big]
\end{aligned}
$$

According to (2.17), this is upper bounded by

$$
\begin{aligned}
\ldots &\overset{(2.17)}{\leq} \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right] \\
&\quad \cdot \Bigg(\sum_{k \leq l_1 \leq s} \cdots \sum_{l_{n-2} \leq l_{n-1} \leq s} E\Big[e^{-\theta\big([S_1 - A_1]^+(k,l_1) + [S_2 - (A_2 + A_1 \oslash S_1)]^+(l_1,l_2) + \ldots} \\
&\quad \cdot e^{+\big[S_n - \big(A_n + A_{n-1} \oslash S_{n-1} + \cdots + \big(\big((A_1 \oslash S_1) \oslash [S_2 - A_2]^+\big) \oslash \cdots \oslash [S_{n-1} - (A_2 + \cdots + A_{n-1})]^+\big)\big)\big]^+ (l_{n-1},s)\big)}\Big]\Bigg)
\end{aligned}
$$

where all factors are linearly dependent, since A_1 occurs in every factor. This requires us to apply the generalized Hölder inequality from Section 2.5:

$$
\begin{aligned}
\ldots &\overset{(2.22)}{\leq} \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right] \left(\sum_{k \leq l_1 \leq s} \cdots \sum_{l_{n-2} \leq l_{n-1} \leq s} E\left[e^{-p_1 \theta [S_1 - A_1]^+(k,l_1)}\right]^{\frac{1}{p_1}}\right. \\
&\quad \left. \cdot E\left[e^{-p_n \theta \big[S_n - \big(A_n + A_{n-1} \oslash S_{n-1} + \cdots + \big(\big((A_1 \oslash S_1) \oslash [S_2 - A_2]^+\big) \oslash \cdots \oslash [S_{n-1} - (A_2 + \cdots + A_{n-1})]^+\big)\big)\big]^+ (l_{n-1},s)}\right]^{\frac{1}{p_n}}\right) \\
&\leq \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right] \left(\sum_{k \leq l_1 \leq s} \cdots \sum_{l_{n-2} \leq l_{n-1} \leq s} E\left[e^{p_1 \theta (A_1(k,l_1) - S_1(k,l_1))}\right]^{\frac{1}{p_1}}\right.
\end{aligned}
$$

$$\cdot E\left[e^{p_n\theta\left(\left(A_n+A_{n-1}\oslash S_{n-1}+\cdots+\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)\right)-S_n\right)(l_{n-1},s)}\right]^{\frac{1}{p_n}}\Bigg)$$

$$= \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right]\left(\sum_{k\le l_1\le s}\cdots\sum_{l_{n-2}\le l_{n-1}\le s} E\left[e^{p_1\theta A_1(k,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(k,l_1)}\right]^{\frac{1}{p_1}}\right.$$

$$\cdot\, E\left[e^{p_n\theta A_n(l_{n-1},s)}\right]^{\frac{1}{p_n}}\cdots E\left[e^{p_n\theta\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)(l_{n-1},s)}\right]^{\frac{1}{p_n}}$$

$$\left.\cdot E\left[e^{-p_n\theta S_n(l_{n-1},s)}\right]^{\frac{1}{p_n}}\right)$$

with

$$\sum_{i=1}^{n}\frac{1}{p_i}=1.$$

Example D.1 (3 servers sink tree output bound with SFA)**.** In the 3-server case, our output bound looks as follows:

$$\phi_{A'(s,t)}(\theta)$$

$$\le \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right]\left(\sum_{k\le l_1\le s}\sum_{l_1\le l_2\le s} E\left[e^{p_1\theta A_1(k,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(k,l_1)}\right]^{\frac{1}{p_1}}\right.$$

$$\cdot\, E\left[e^{p_2\theta A_2(l_1,l_2)}\right]^{\frac{1}{p_2}} E\left[e^{p_2\theta(A_1\oslash S_1)(l_1,l_2)}\right]^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,l_2)}\right]^{\frac{1}{p_2}}$$

$$\left.\cdot E\left[e^{p_3\theta A_3(l_2,s)}\right]^{\frac{1}{p_3}} E\left[e^{p_3\theta\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)(l_2,s)}\right]^{\frac{1}{p_3}} E\left[e^{-p_3\theta S_3(l_2,s)}\right]^{\frac{1}{p_3}}\right)$$

$$\le \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right]\left(\sum_{k\le l_1\le s}\sum_{l_1\le l_2\le s} E\left[e^{p_1\theta A_1(k,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(k,l_1)}\right]^{\frac{1}{p_1}}\right.$$

$$\cdot\, E\left[e^{p_2\theta A_2(l_1,l_2)}\right]^{\frac{1}{p_2}}\left(\sum_{r=0}^{l_1} E\left[e^{p_2\theta(A_1(r,l_2)-S_1(r,l_1))}\right]\right)^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,l_2)}\right]^{\frac{1}{p_2}}$$

$$\left.\cdot E\left[e^{p_3\theta A_3(l_2,s)}\right]^{\frac{1}{p_3}}\left(\sum_{u=0}^{l_2} E\left[e^{p_3\theta\left((A_1\oslash S_1)(u,s)-[S_2-A_2]^+(u,l_2)\right)}\right]\right)^{\frac{1}{p_3}} E\left[e^{-p_3\theta S_3(l_2,s)}\right]^{\frac{1}{p_3}}\right)$$

$$\overset{(2.21)}{\le} \sum_{k=0}^{s} E\left[e^{\theta A_{\mathrm{foi}}(k,t)}\right]\left(\sum_{k\le l_1\le s}\sum_{l_1\le l_2\le s} E\left[e^{p_1\theta A_1(k,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(k,l_1)}\right]^{\frac{1}{p_1}}\right.$$

$$\cdot\, E\left[e^{p_2\theta A_2(l_1,l_2)}\right]^{\frac{1}{p_2}}\left(\sum_{r=0}^{l_1} E\left[e^{p_2\theta A_1(r,l_2)}\right] E\left[e^{-p_2\theta S_1(r,l_1)}\right]\right)^{\frac{1}{p_2}} E\left[e^{-p_2\theta S_2(l_1,l_2)}\right]^{\frac{1}{p_2}}$$

$$\cdot E\left[e^{p_3\theta A_3(l_2,s)}\right]^{\frac{1}{p_3}} \left(\sum_{u=0}^{l_2}\left(\sum_{v=0}^{u} E\left[e^{p_3\theta A_1(v,s)}\right] E\left[e^{-p_3\theta S_1(v,u)}\right]\right) E\left[e^{p_3\theta A_2(u,l_2)}\right] E\left[e^{-p_3\theta S_2(u,l_2)}\right]\right)^{\frac{1}{p_3}}$$
$$\cdot E\left[e^{-p_3\theta S_3(l_2,s)}\right]^{\frac{1}{p_3}}\Bigg)$$

and

$$\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}=1.$$

D.2. Proposition 7.3

$$\begin{aligned}
\mathrm{P}(d(t)>T) &\overset{(2.13)}{\leq} E\left[e^{\theta(A_{\text{foi}}\oslash S_{\text{l.o.}}(t+T,t))}\right]\\
&\overset{(2.7)}{=} E\left[e^{\theta\max_{0\leq s\leq t}\{A_{\text{foi}}(s,t)-S_{\text{l.o.}}(s,t+T)\}}\right]\\
&\overset{(2.21)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta(A_{\text{foi}}(s,t)-S_{\text{l.o.}}(s,t+T))}\right]\\
&= \sum_{s=0}^{t} E\left[e^{\theta A_{\text{foi}}(s,t)}\right] E\Big[e^{-\theta\big([S_1-A_1]^+\otimes[S_2-(A_2+A_1\oslash S_1)]^+\otimes}\\
&\quad\cdot\ {}^{\cdots\otimes\left[S_n-\left(A_n+A_{n-1}\oslash S_{n-1}+\cdots+\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)\right)\right]^+(s,t+T)\big)}\Big].
\end{aligned}$$

According to (2.17), this is upper bounded by

$$\begin{aligned}
\ldots &\overset{(2.17)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta A_{\text{foi}}(s,t)}\right]\Bigg(\sum_{s\leq l_1\leq t+T}\cdots\sum_{l_{n-2}\leq l_{n-1}\leq t+T} E\Big[e^{-\theta\big([S_1-A_1]^+(s,l_1)+[S_2-(A_2+A_1\oslash S_1)]^+(l_1,l_2)+\ldots}\\
&\quad\cdot e^{+\left[S_n-\left(A_n+A_{n-1}\oslash S_{n-1}+\cdots+\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)\right)\right]^+(l_{n-1},t+T)\big)}\Big]\Bigg).
\end{aligned}$$

By the generalized Hölder inequality (2.22) (because of the factors' dependency), we get

$$\begin{aligned}
\ldots &\overset{(2.22)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta A_{\text{foi}}(s,t)}\right]\Bigg(\sum_{s\leq l_1\leq t+T}\cdots\sum_{l_{n-2}\leq l_{n-1}\leq t+T} E\left[e^{-p_1\theta[S_1-A_1]^+(s,l_1)}\right]^{\frac{1}{p_1}}\\
&\quad\cdots E\left[e^{-p_n\theta\left[S_n-\left(A_n+A_{n-1}\oslash S_{n-1}+\cdots+\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)\right)\right]^+(l_{n-1},t+T)}\right]^{\frac{1}{p_n}}\Bigg)\\
&\leq \sum_{s=0}^{t} E\left[e^{\theta A_{\text{foi}}(s,t)}\right]\Bigg(\sum_{s\leq l_1\leq t+T}\cdots\sum_{l_{n-2}\leq l_{n-1}\leq t+T} E\left[e^{p_1\theta(A_1-S_1)(s,l_1)}\right]^{\frac{1}{p_1}}
\end{aligned}$$

$$\cdots E\left[e^{p_n\theta\left(\left(A_n+A_{n-1}\oslash S_{n-1}+\cdots+\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)\right)-S_n\right)(l_{n-1},t+T)}\right]^{\frac{1}{p_n}}\right)$$

$$= \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right]\left(\sum_{s\le l_1\le t+T}\cdots\sum_{l_{n-2}\le l_{n-1}\le t+T} E\left[e^{p_1\theta A_1(s,l_1)}\right]^{\frac{1}{p_1}} E\left[e^{-p_1\theta S_1(s,l_1)}\right]^{\frac{1}{p_1}}\right.$$

$$\cdots E\left[e^{p_n\theta\left(A_n+A_{n-1}\oslash S_{n-1}+\cdots+\left(\left((A_1\oslash S_1)\oslash[S_2-A_2]^+\right)\oslash\cdots\oslash[S_{n-1}-(A_2+\cdots+A_{n-1})]^+\right)\right)(l_{n-1},t+T)}\right]^{\frac{1}{p_n}}$$

$$\left.\cdot E\left[e^{-p_n\theta S_n(l_{n-1},t+T)}\right]^{\frac{1}{p_n}}\right)$$

with

$$\sum_{i=1}^{n}\frac{1}{p_i}=1.$$

D.3. Proposition 7.7

$$\begin{aligned}
\mathrm{P}(d(t)>T) &\overset{(2.13)}{\le} E\left[e^{\theta(A_{\mathrm{foi}}\oslash S_{\mathrm{l.o.}}(t+T,t))}\right]\\
&\overset{(2.7)}{=} E\left[e^{\theta\max_{0\le s\le t}\{A_{\mathrm{foi}}(s,t)-S_{\mathrm{l.o.}}(s,t+T)\}}\right]\\
&\overset{(2.21)}{\le} \sum_{s=0}^{t} E\left[e^{\theta(A_{\mathrm{foi}}(s,t)-S_{\mathrm{l.o.}}(s,t+T))}\right]\\
&= \sum_{s=0}^{t} E\left[e^{\theta\left(A_{\mathrm{foi}}(s,t)-\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_1\right)-A_1\right]^+(s,t+T)\right)}\right]\\
&= \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{-\theta\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_1\right)-A_1\right]^+(s,t+T)}\right]\\
&\le \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{\theta\left(A_1(s,t+T)-\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_1\right)(s,t+T)\right)}\right]\\
&= \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{\theta A_1(s,t+T)}\right]\\
&\quad\cdot E\left[e^{-\theta\left(\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots S_2\right)-A_2\right]^+\otimes S_1\right)(s,t+T)}\right]\\
&\overset{(2.21)}{\le} \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{\theta A_1(s,t+T)}\right]
\end{aligned}$$

$$
\begin{aligned}
&\cdot \sum_{l_1=s}^{t+T} E\left[e^{\theta\left(-\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots S_2\right)-A_2\right]^+(s,l_1)-S_1(l_1,t+T)\right)}\right] \\
&= \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{\theta A_1(s,t+T)}\right] \\
&\cdot \sum_{l_1=s}^{t+T} E\left[e^{-\theta S_1(l_1,t+T)}\right] E\left[e^{-\theta\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots S_2\right)-A_2\right]^+(s,l_1)}\right] \\
&\leq \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{\theta A_1(s,t+T)}\right] \sum_{l_1=s}^{t+T} E\left[e^{-\theta S_1(l_1,t+T)}\right] E\left[e^{\theta A_2(s,l_1)}\right] \\
&\cdot \sum_{l_2=s}^{l_1} E\left[e^{-\theta\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots S_3\right)-A_3\right]^+(s,l_2)}\right] E\left[e^{-\theta S_2(l_2,l_1)}\right] \\
&\vdots \\
&\leq \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] \mathbb{E}\left[e^{\theta A_1(s,t+T)}\right] \left(\sum_{l_1=s}^{t+T} E\left[e^{-\theta S_1(l_1,t+T)}\right] E\left[e^{\theta A_2(s,l_1)}\right]\right. \\
&\cdot \left(\sum_{l_2=s}^{l_1} E\left[e^{-\theta S_2(l_2,l_1)}\right] E\left[e^{\theta A_3(s,l_2)}\right] \cdots \left(\sum_{l_k=s}^{l_{k-1}} E\left[e^{-\theta S_k(l_k,l_{k-1})}\right] E\left[e^{\theta A_{k+1}(s,l_k)}\right]\right.\right. \\
&\left.\left.\left.\cdots \left(\sum_{l_{n-1}=s}^{l_{n-2}} E\left[e^{-\theta S_{n-1}(l_{n-1},l_{n-2})}\right] E\left[e^{\theta A_n(s,l_{n-1})}\right] E\left[e^{-\theta S_n(s,l_{n-1})}\right]\right)\right)\right)\right).
\end{aligned}
$$

D.4. Proposition 7.10

$$
\begin{aligned}
\mathrm{P}(d(t)>T) &\overset{(2.13)}{\leq} E\left[e^{\theta(A_{\mathrm{foi}}\oslash S_{\mathrm{l.o.}}(t+T,t))}\right] \\
&= E\left[e^{\theta \max_{0\leq s\leq t}\{A_{\mathrm{foi}}(s,t)-S_{\mathrm{l.o.}}(s,t+T)\}}\right] \\
&\overset{(2.21)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta(A_{\mathrm{foi}}(s,t)-S_{\mathrm{l.o.}}(s,t+T))}\right] \\
&= \sum_{s=0}^{t} E\left[e^{\theta\left(A_{\mathrm{foi}}(s,t)-\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_1\right)-A_1\right]^+(s,t+T)\right)}\right] \\
&= \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{-\theta\left[\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_1\right)-A_1\right]^+(s,t+T)}\right] \\
&\leq \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{\theta\left(A_1(s,t+T)-\left(\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_1\right)(s,t+T)\right)}\right].
\end{aligned}
$$

Here, we have A_1 being dependent on the other A_i's and hence, Hölder's inequality comes into play.

$$
\begin{aligned}
\mathrm{P}(d(t)>T) &\leq \ldots \\
&\overset{(2.20)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] E\left[e^{p_1\theta A_1(s,t+T)}\right]^{\frac{1}{p_1}} \\
&\quad\cdot \left(E\left[e^{-p_2\theta\left(\left[\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots-A_2\right]^+\otimes S_1\right)(s,t+T)}\right]\right)^{\frac{1}{p_2}} \\
&\overset{(2.21)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] \left(E\left[e^{p_1\theta A_1(s,t+T)}\right]\right)^{\frac{1}{p_1}} \\
&\quad\cdot \left(\sum_{l_1=s}^{t+T} E\left[e^{-p_2\theta\left[\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots-A_2\right]^+(s,l_1)}\right] E\left[e^{-p_2\theta S_1(l_1,t+T)}\right]\right)^{\frac{1}{p_2}} \\
&\leq \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] \left(E\left[e^{p_1\theta A_1(s,t+T)}\right]\right)^{\frac{1}{p_1}} \\
&\quad\cdot \left(\sum_{l_1=s}^{t+T} E\left[e^{-p_2\theta S_1(l_1,t+T)}\right] E\left[e^{p_2\theta\left(A_2-\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_2\right)(s,l_1)}\right]\right)^{\frac{1}{p_2}}
\end{aligned}
$$

and we see that Hölder has to be applied n-times:

$$
\begin{aligned}
\mathrm{P}(d(t)>T) &\leq \ldots \\
&\overset{(2.20)}{\leq} \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] \left(E\left[e^{p_1\theta A_1(s,t+T)}\right]\right)^{\frac{1}{p_1}} \\
&\quad\cdot \left(\sum_{l_1=s}^{t+T} E\left[e^{-p_2\theta S_1(l_1,t+T)}\right] \left(E\left[e^{p_2p_3\theta A_2(s,l_1)}\right]\right)^{\frac{1}{p_3}}\right. \\
&\quad\cdot \left.\left(E\left[e^{-p_2p_4\theta\left[\left([S_n-A_n]^+\otimes S_{n-1}\right)-A_{n-1}\right]^+\otimes\cdots\otimes S_2(s,l_1)}\right]\right)^{\frac{1}{p_4}}\right)^{\frac{1}{p_2}} \\
&\quad\vdots \\
&\leq \sum_{s=0}^{t} E\left[e^{\theta A_{\mathrm{foi}}(s,t)}\right] \left(E\left[e^{p_1\theta A_1(s,t+T)}\right]\right)^{\frac{1}{p_1}} \\
&\quad\cdot \left(\sum_{l_1=s}^{t+T} E\left[e^{-p_2\theta S_1(l_1,t+T)}\right] \left(E\left[e^{p_2p_3\theta A_2(s,l_1)}\right]\right)^{\frac{1}{p_3}}\right.
\end{aligned}
$$

$$
\begin{aligned}
&\cdots\left(\sum_{l_{k-1}=s}^{l_{k-2}} E\left[e^{-p_2p_4\cdots p_{2k-2}\theta S_{k-1}(l_{k-1},l_{k-2})}\right]\left(E\left[e^{p_2p_4\cdots p_{2k-2}p_{2k-1}\theta A_k(s,l_{k-1})}\right]\right)^{\frac{1}{p_{2k-1}}}\right.\\
&\cdots\left(\sum_{l_{n-2}=s}^{l_{n-3}} E\left[e^{-p_2p_4\cdots p_{2n-4}\theta S_{n-2}(l_{n-2},l_{n-3})}\right]\left(E\left[e^{p_2p_4\cdots p_{2n-4}p_{2n-3}\theta A_{n-1}(s,l_{n-1})}\right]\right)^{\frac{1}{p_{2n-3}}}\right.\\
&\cdot\left(\sum_{l_{n-1}=s}^{l_{n-2}} E\left[e^{-p_2p_4\cdots p_{2n-2}\theta S_{n-1}(l_{n-1},l_{n-2})}\right]E\left[e^{p_2p_4\cdots p_{2n-2}\theta A_n(s,l_{n-1})}\right]\right.\\
&\left.\cdot\ E\left[e^{-p_2p_4\cdots p_{2n-2}\theta S_n(s,l_{n-1})}\right]\right)^{\frac{1}{p_{2n-2}}}\\
&\left.\left.\left.\cdots\right)^{\frac{1}{p_{2k-2}}}\right)^{\frac{1}{p_4}}\right)^{\frac{1}{p_2}}
\end{aligned}
$$

with

$$
\begin{aligned}
\frac{1}{p_1}+\frac{1}{p_2}&=1\\
\vdots&=\vdots\\
\frac{1}{p_{2n-1}}+\frac{1}{p_{2n-2}}&=1.
\end{aligned}
$$

E. Equivalence of the Max-Sum and the Union Bound

The max-sum bound (2.21) from Section 2.5 and the union bound are closely related, as we see in this chapter. At first, let us restate the union bound for this purpose.

Proposition E.1 (Union bound)**.** *For a countably many sets $A_1, A_2, \ldots,$ we have*

$$\mathrm{P}\left(\bigcup_i A_i\right) \leq \sum_i \mathrm{P}(A_i)\,. \tag{E.1}$$

Next, we show that the max-sum bound (2.21) is equivalent to the union bound (E.1) if the Chernoff inequality is applied (this result can also be found in [Bec16]).

Theorem E.2 (Interchangeability of max-sum and union bound)**.** *Let X_i, $i = 1, \ldots, n$ be integrable random variables. Whether we bound $\mathrm{P}(\max_{1\leq i\leq n} X_i > x)$ by the union bound and then apply the Chernoff bound or use the Chernoff bound at first and the max-sum bound afterwards, we obtain the same result, i.e., the order does not influence the result.*

Proof. Case I: we apply the union bound at first and the Chernoff bound afterwards:

$$\begin{aligned}
\mathrm{P}\left(\max_{1\leq i\leq n}\{X_i > x\}\right) &= \mathrm{P}\left(\bigcup_{i=1}^{n} X_i > x\right) \\
&\overset{(\mathrm{E.1})}{\leq} \sum_{i=1}^{n} \mathrm{P}(X_i > x) \\
&\overset{(2.2)}{\leq} \sum_{i=1}^{n} e^{-\theta x}\,\mathrm{E}\Big[e^{\theta X_i}\Big]\,,
\end{aligned}$$

compared to case II, applying the Chernoff bound and then the max-sum bound:

$$\begin{aligned}
\mathrm{P}\left(\max_{1\leq i\leq n}\{X_i > x\}\right) &\overset{(2.2)}{\leq} e^{-\theta x}\,\mathrm{E}\Big[e^{\theta \max_{1\leq i\leq n}\{X_i\}}\Big] \\
&\overset{(2.21)}{\leq} \sum_{i=1}^{n} e^{-\theta x}\,\mathrm{E}\Big[e^{\theta X_i}\Big]
\end{aligned}$$

and the proof is complete. □

F. Additional Results (Lyapunov Inequality)

This chapter summarizes several mathematical results that are made use of in this thesis.

F.1. The Root Function is Concave

Proposition F.1 (Concavity of the root function)**.** *The p-th root function*

$$\begin{aligned} f : \mathbb{R}_0^+ &\to \mathbb{R}_0^+ \\ x &\mapsto x^{\frac{1}{p}} \end{aligned}$$

is concave for $p \geq 1$.

Proof. For $x > 0$ and $p > 1$, we have strict concavity since

$$\begin{aligned} f'(x) &= \frac{1}{p} x^{\frac{1}{p}-1}, \\ f''(x) &= \frac{1}{p} \underbrace{\left(\frac{1}{p} - 1\right)}_{<0} \underbrace{x^{\frac{1}{p}-2}}_{>0} < 0 \end{aligned}$$

for all $x \geq 0$ and for $p = 1$, it holds for $\alpha \in [0, 1]$ that

$$f\left((1-\alpha)\, x + \alpha y\right) = (1-\alpha)\, x + \alpha y = (1-\alpha)\, f(x) + \alpha f\,(y)$$

and thus the desired " $\geq$ " relation holds with equality. At last, for $x = 0$ and $p \geq 1$, we have

$$f\left((1-\alpha)\, x + \alpha y\right) = f\left(\alpha y\right) = \alpha^{\frac{1}{p}} y^{\frac{1}{p}} \overset{(\alpha \leq 1)}{\geq} \alpha y^{\frac{1}{p}} = \alpha f\,(y) = (1-\alpha)\, f(x) + \alpha f\,(y)$$

and so, the proof is complete. □

F.2. Jensen's Inequality

Theorem F.2 (Jensen's inequality)**.** *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex and let X and $\varphi(x)$ be integrable random variables. For any σ-algebra $\mathcal{G}$,*

$$\varphi\left(\mathrm{E}[X \mid \mathcal{G}]\right) \leq \mathrm{E}[\varphi(x) \mid \mathcal{G}]\,. \tag{F.1}$$

If ψ is concave and $\psi(x)$ is integrable, then it holds vice versa

$$\psi\left(\mathrm{E}[X \mid \mathcal{G}]\right) \geq \mathrm{E}[\psi(x) \mid \mathcal{G}]\,. \tag{F.2}$$

Proof. See, e.g., [JP02]. □

F.3. Concavity and Subadditivity

Proposition F.3. *Let $f : \mathbb{R}_0^+ \to \mathbb{R}$ be concave and $f(0) \geq 0$. Then f is also subadditive.*

Proof. For $x, y \geq 0$ and $a \in [0,1]$, it holds that

$$f\left(ax\right) = f\left(ax + (1-a)\cdot 0\right) \overset{\text{concave}}{\geq} af(x) + \underbrace{(1-a)}_{\geq 0} \cdot \underbrace{f(0)}_{\geq 0} \geq af(x). \tag{F.3}$$

Then

$$\begin{aligned} f(x) + f\left(y\right) &= f\left(\left(x+y\right)\frac{x}{x+y}\right) + f\left(\left(x+y\right)\frac{y}{x+y}\right) \\ &\overset{\text{(F.3)}}{\geq} \frac{x}{x+y} f\left(x+y\right) + \frac{y}{x+y} f\left(x+y\right) \\ &= f\left(x+y\right), \end{aligned}$$

since $\frac{x}{x+y}, \frac{y}{x+y} \in [0,1]$ (like a). This finishes the proof. □

Remark F.4. This result can also be found in [LT01].

G. Markov Chains

The following definitions and theorems can be found in [KKK10].

Definition G.1 (Markov chain)**.** Let $\{X_t, t \in \mathbb{N}\}$ be a discrete-time stochastic process such that X_t only attains values in a countable set S, the so-called *state space*. For convenience we always identify S with (a subset of) $\mathbb{N}$. It is called a (discrete-time) *Markov chain*, if we have

$$\begin{aligned} \mathrm{P}(X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, \dots X_0 = i_0) \\ = \mathrm{P}(X_{t+1} = j \mid X_t = i) =: p_{i,j}(t) \end{aligned}$$

for all $i, j, i_k \in S$. The matrix $p_{i,j}(t)$ is called the transition matrix from time t to time $t+1$.

Definition G.2 (Homogeneous and recurrent)**.** A Markov chain $\{X_t\}_{t \in \mathbb{N}}$ with state space S is called *homogeneous* if for any $i, j \in S$ and any $n \in \mathbb{N}$ we have

$$\mathrm{P}(X_{t+1} = j \mid X_t = i) = \mathrm{P}(X_1 = j \mid X_0 = i) = p_{i,j}. \tag{G.1}$$

Let $i \in S$ be some state of the homogeneous Markov chain $\{X_t\}_{t \in \mathbb{N}}$ and $X_0 = i$. Denote by

$$\tau_i = \min \{t > 0 \mid X_t = i\}$$

the first recurrence time. The state i is called *recurrent*, if we have

$$\mathrm{P}(\tau_i < \infty) = 1.$$

If the recurrent state i satisfies

$$\mathrm{E}[\tau_i] < \infty$$

then it is called *positive recurrent*.

Definition G.3 (Connected and aperiodic)**.** Two states $i, j \in S$ of the homogeneous Markov chain $\{X_t\}_{t \in \mathbb{N}}$ are called *connected*, if for some $t_{12}, t_{21} \in \mathbb{N}$ we have

$$\mathrm{P}(X_{t+t_{12}} = j \mid X_t = i) \cdot \mathrm{P}(X_{t+t_{21}} = i \mid X_t = j) > 0.$$

The Markov chain is called *aperiodic* if for all states $i \in S$ we have

$$1 = \gcd \{t \in \mathbb{N} \mid \mathrm{P}(X_t = i \mid X_0 = i) > 0\}.$$

Definition G.4 (Irreducible)**.** A homogeneous Markov chain $\{X_t\}_{t \in \mathbb{N}}$ with state space S is called *irreducible* if it has only one equivalence class of connected states.

Definition G.5 (Stationary distribution)**.** A distribution $\pi(.)$ of the state S is called a *stationary distribution* of the homogeneous Markov chain $\{X_t\}_{t\in\mathbb{N}}$ if all X_t are distributed according to $\pi(.)$ when the starting value X_0 is distributed according to $\pi(.)$.

Remark G.6. Usually, a stationary distribution π of a Markov chain is defined as a nonnegative solution of the equations

$$\pi p = \pi, \sum \pi(i) = 1. \tag{G.2}$$

The Theorem below shows that the two definitions are essentially equivalent.

Theorem G.7. *Let $\{X_t\}_{t\in\mathbb{N}}$ be a homogeneous, irreducible, and aperiodic Markov chain with positive recurrent states. Then it has a unique stationary distribution π that satisfies*

$$\lim_{t\to\infty} p_{i,j}^{(t)} = \pi(j) \tag{G.3}$$

for all $i, j \in S$ and is the unique nonnegative solution to equation (G.2).

H. System of Linear Differential Equations

The following results are explained in detail in [Heu13] (in particular in Chapter 49).

Proposition H.1. *The solution to*

$$\begin{aligned} X'(t) &= AX(t) \\ X(0) &= I, \end{aligned}$$

where I denotes the identity matrix, is

$$\exp(At) \coloneqq I + At + \frac{A^2t^2}{2} \dots$$

Proposition H.2. *Let A be diagonalizable, i.e., $A = Q^{-1}\Omega Q$, $\det(Q) \neq 0$ and $\Omega = diag(\omega_1, \dots, \omega_n)$. Then, it holds that*

$$\exp(A) = Q^{-1} diag(e^{\omega_1}, \dots, e^{\omega_n}) Q.$$

For the 2-dimensional case, this means

$$\exp(At) = Q^{-1} \begin{pmatrix} e^{\omega_1 t} & 0 \\ 0 & e^{\omega_2 t} \end{pmatrix} Q.$$

I. Fractional Brownian Motion

We repeat Definition 2.1 in [MVN68] and restate the most important properties of fractional Brownian motion.

Definition I.1 (Fractional Brownian motion)**.** Let $H \in (0,1)$ and let $f_0 \in \mathbb{R}$ be arbitrary. We call the following stochastic process $F_H(t,\omega)$ *fractional Brownian motion* (fBm) with parameter H and starting value f_0 at time 0. For $t > 0$, $F_H(t,\omega)$ is defined by

$$\begin{aligned} F_H(0,\omega) =& f_0, \\ F_H(t,\omega) - F_H(0,\omega) =& \frac{1}{\Gamma\left(H+\frac{1}{2}\right)} \left(\int_{-\infty}^{0} (t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \mathrm{d}F \right. \\ & \left. + \int_0^t (t-s)^{H-\frac{1}{2}} \mathrm{d}F(s,\omega) \right), \end{aligned} \tag{I.1}$$

with Gamma function

$$\Gamma(t) := \int_0^\infty x^{t-1} e^{-x} \mathrm{d}x. \tag{I.2}$$

The increment process

$$\Delta F_H(t,\omega) = F_H(t,\omega) - F_H\left(t-1,\omega\right)$$

is called *fraction Gaussian noise* (fGn).

Definition I.2 (Self-similarity)**.** The increments of a stochastic process $X(t,\omega)$ are said to be self-similar with parameter $H \geq 0$ if for any $a > 0$ and any t_0

$$X\left(t_0 + a\tau, \omega\right) - X\left(t_0, \omega\right)$$

and

$$a^H \left(X\left(t_0 + \tau, \omega\right) - X\left(t_0, \omega\right) \right)$$

have the same distribution.

Theorem I.3. *The increments of fBm, i.e., fGn, are stationary and self-similar with parameter H.*

Proof. See Theorem 3.3 in [MVN68]. □