# TOWARDS A STATISTICAL NETWORK CALCULUS– DEALING WITH UNCERTAINTY IN ARRIVALS

MICHAEL A. BECK, SEBASTIAN A. HENNINGSEN, SIMON B. BIRNBACH, JENS B. SCHMITT
DISTRIBUTED COMPUTER SYSTEMS (DISCO) LAB, UNIVERSITY OF KAISERSLAUTERN, GERMANY

Abstract. The stochastic network calculus (SNC) has become an attractive methodology to derive probabilistic performance bounds. So far the SNC is based on (tacitly assumed) exact probabilistic assumptions about the arrival processes. Yet, in practice, these are only true approximately–at best. In many situations it is hard, if possible at all to make such assumptions a priori. A more practical approach would be to base the SNC operations on measurements of the arrival processes (preferably even on-line). In this report, we develop this idea and incorporate measurements into the framework of SNC taking the further uncertainty resulting from estimation errors into account. This is a crucial step towards a statistical network calculus (StatNC) eventually lending itself to a self-modelling operation of networks with a minimum of a priori assumptions. In numerical experiments, we are able to substantiate the novel opportunities by StatNC.

## 1. Introduction

1.1. **Motivation.** Over the last two decades the stochastic network calculus (SNC) has evolved as a valuable methodology to compute probabilistic performance bounds [8]. It has found numerous and diverse usage in important network design and control problems: smart grid control [25], delay control in cognitive radio networks [13], and as foundation for bandwidth estimation on Internet end-to-end paths [21], to name a few recent examples.

SNC originated from its deterministic counterpart as conceived by Cruz [10, 11] to provide stochastically relaxed performance bounds, mainly in order to capture the statistical multiplexing gain as is characteristic for packet-switched networks. Some of the earliest work on SNC can be traced back to [26, 5, 19]. In particular Chang's sigma-rho calculus based on moment-generating functions (MGF) received much attraction in the field and was refined in [12] to match with the latest advances in the min-plus algebraic formulation of network calculus (alternative SNC formulations can be found in [12, 4, 18], see [8] for some perspectives about these). The core modelling abstractions of SNC are arrival envelopes and service curves. Arrival envelopes provide probabilistic bounds on how much traffic arrives within a time interval of a given length; service curves essentially do the same for the amount of work done by a system serving those arrivals.

One of the strengths of SNC is its versatility with respect to traffic models that can be treated, ranging from short-range dependent traffic with exponentially bounded burstiness (see e.g. [7]) to long-range dependent traffic such as fractional Brownian motion [24], or even heavy-tailed self-similar traffic [20]. Yet, all of these works start from "clean", a priori and exact probabilistic assumptions. In practice, however, the question arises: where do these assumptions come from? In most cases the answer must be: observation of the past traffic behaviour, in the form of measurements and subsequent statistical inference. However, statistical inference involves errors and thus another source of uncertainty besides traffic variations themselves. To the best of our knowledge, none of the existing work on SNC has taken this uncertainty into account and integrated it into the SNC operations. We take this missing first step, i.e., measuring the arrivals and making statistical inferences, and integrate it into the SNC, thereby moving towards a statistical network calculus (StatNC)[1]. Moving from SNC to StatNC can be viewed as going from stochastic processes to time series. Of course, we still have to make assumptions for the time series corresponding to past traffic arrivals with respect to the underlying stochastic process, but we can adapt them dynamically (possibly on-line) and some deviations from the assumptions may be tolerable (depending on the robustness of our statistical estimators). Clearly, the goal of our StatNC framework is to cope with as few assumptions as possible while still providing accurate performance bounds.

To illustrate where a statistical network calculus can be very beneficial, let us briefly sketch two application scenarios:

(1) **Traffic engineering in an MPLS domain** [2]: traffic is measured at ingress nodes to an MPLS domain and label-switched paths are dynamically dimensioned according to service level agreements based on StatNC; an immediate benefit is that time-of-day effects or any other seasonal effects are automatically taken into account.
(2) **Self-modelling in wireless sensor networks**: traffic is measured at sensor nodes and the resulting estimates are delivered towards a sink (in the simplest case) which can then base decisions such as, e.g., topology control on the respective StatNC models; an immediate benefit is that no a priori traffic description is necessary any more, which is very helpful in many WSN applications as the behaviour of the physical phenomena to be observed is often not well-understood before deployment and thus the traffic induced by them is hard to predict.

Overall, we make the following contributions

- development of a uniform framework for a statistical network calculus which allows to plug in a large class of traffic estimators ($\rightarrow$ Section 3);
- design of several traffic estimators with differing amount of presumed knowledge and probabilistic assumptions ($\rightarrow$ Section 4);
- in numerical examples, the practicality, precision, and robustness of the StatNC is investigated and contrasted against the performance of SNC alongside with simulative results ($\rightarrow$ Section 5).

1.2. **Related Work.** In the SNC literature, there are only few papers that discuss the fact that arrival envelopes could be derived from measurements: for example, [8] provides a brief sketch how a measured packet trace could be fitted to a weighted

[1]The term statistical network calculus has been used before to indicate that the SNC takes into account statistical multiplexing gains [4], whereas here we use it to indicate the usage of statistical methods instead of purely probabilistic reasoning.

hyperexponential traffic envelope, while [20] even does it for a heavy-tailed self-similar traffic envelope. Yet, none of these integrates the measurements with the SNC operations such that the uncertainty resulting from estimation errors is factored into the stochastic bounds. In the report at hand, we perform this integration in a rigorous and uniform manner (see Theorem 5 in Section 3). Furthermore, to the best of our knowledge there is no previous work in the SNC literature about an on-line estimation of the arrival envelope as it is enabled by our StatNC framework.

In a larger context, somewhat related work can be found in the domain of measurement-based admission control (see e.g. [17, 15, 23, 14]]). However, because at that time the SNC was not yet fully developed, these works are restricted to admission control rather than basing on a general performance evaluation framework like SNC. Furthermore, they typically assume a known (deterministic) traffic envelope and then measure to what extent this envelope is used and how statistical multiplexing helps to reduce resource demands, whereas in our work we basically start one step earlier by estimating the probabilistic arrival envelopes themselves.

Also slightly related is the work by Lübben et al. [21] on the identification of stochastic service curves to represent Internet end-to-end paths. Clearly, measurements (though active ones) play a central role here as well, yet the target is different in our case as we deal with the uncertainty about arrival rather than service processes.

On a very high level, the vision of autonomic networking (see e.g. [3] for a prominent large-scale project in that domain) could be related especially to the self-modelling aspect of the StatNC when used in an on-line fashion, yet no use of SNC within this domain is known to us, although it appears to be a very promising idea.

## 2. Preliminaries on Stochastic Network Calculus

In this report, we focus on the SNC formulation as originally presented in [6] and later on generalized by [12], which is also known as $(\sigma(\theta), \rho(\theta))$-calculus. In this setup, time is discrete while data is allowed to be continuous (i.e., we deal with infinitesimally small data units). For convenience, we make a few small modifications to definitions and notations from [6], and therefore repeat the most important of them together with the main results needed in this report. Since, for brevity, we focus on the backlog as performance measure in this report, we only present the corresponding results. More results, concerning other performance measures (i.e., virtual delay and output bounds) and about reducing the complexity of networks with multiple flows and service elements, can be derived in a similar fashion.

In SNC, data flows arrive at service elements and after processing leave them again. To represent such flows, we define a real non-negative stochastic process $(a_k)_{k\in\mathbb{Z}}$ and the bivariate cumulatives

$$A(m,n) := \sum_{k=m+1}^{n} a_k.$$

We henceforth call the random variables $a_k$ *increments* of the *flow* $A$. Since the basic idea of StatNC is to apply statistical methods on past observations, we think of increments with time index $k < 0$ as lying in the past (the so-far observed time series of arrivals). The increments with index $k \geq 0$ are upcoming arrivals. Performance bounds are always calculated for points in time lying on the positive time axis. Further, we assume a value $n_0 \leq 0$ such that $a_k = 0$ for all $k < n_0$, this is the time when we started our observations.

The service element is also abstracted by a doubly indexed stochastic process $S$ with the properties:

$$\begin{aligned} 0 &\leq S(m,n) \qquad \forall\, m,n \in \mathbb{N}_0 \\ S(m,n) &\leq S(m,n') \qquad \forall\, m,n,n' \in \mathbb{N}_0 \text{ and } n \leq n' \end{aligned}$$

Note that we define $S$ only on $\mathbb{N}_0 \times \mathbb{N}_0$, which is–as we will see–sufficient. The service process $S$, arrival flow $A$ and the departure flow $D$ of a service element are linked with each other in the following way:

**Definition 1.** If for all $n \in \mathbb{N}_0$ holds

$$D(0,n) \geq \min_{0 \leq k \leq n} \{A(0,k) + S(k,n)\},$$

we call the service element a dynamic $S$-server. Here $D$ is defined as a flow with $n_0 = 0$.

Before we can give stochastic bounds on the backlog of a system, we need some bounds on the arrivals and the dynamic $S$-server. More precisely, we need bounds on the moment generating functions (MGF) of the corresponding stochastic processes.

**Definition 2.** Let $\theta > 0$. An arrival is $(\sigma_A(\theta), \rho_A(\theta))$-bounded if

$$\sup_{m \in \mathbb{Z}} \{\mathbb{E}({}^{\theta A(m,m+k)})\} \leq e^{k\theta\rho_A(\theta) + \theta\sigma_A(\theta)} \quad \forall\, k \in \mathbb{N}$$

A dynamic $S$-server is $(\sigma_S(\theta), \rho_S(\theta))$-bounded if

$$\sup_{m \geq 0} \{\mathbb{E}(e^{-\theta S(m,m+k)})\} \leq e^{k\theta\rho_S(\theta) + \theta\sigma_S(\theta)} \quad \forall\, k \in \mathbb{N}$$

We are now able to give stochastic bounds on a service element's *backlog* process defined by $q(n) := A(0,n) - D(0,n)$.

**Theorem 3.** *Let $A$ be an arrival flow served by a dynamic $S$-server and $\theta > 0$. Assume $A$ is $(\sigma_A(\theta), \rho_A(\theta))$-bounded and $S$ is $(\sigma_S(\theta), \rho_S(\theta))$-bounded. If $A$ is stochastically independent of $S$, the following probabilistic bound holds:*

$$\mathbb{P}(q(n) > x) \leq e^{-\theta x} e^{\theta(\sigma_A(\theta) + \sigma_S(\theta))} \sum_{k=0}^{n} e^{k\theta(\rho_A(\theta) + \rho_S(\theta))}$$

*If $A$ is not stochastically independent of $S$ we still have:*

$$\mathbb{P}(q(n) > x) \leq e^{-\theta x} e^{\theta(\sigma_A(p\theta) + \sigma_S(q\theta))} \sum_{k=0}^{n} e^{k\theta(\rho_A(p\theta) + \rho_S(q\theta))}$$

*for some $p$ and $q$ such that $p^{-1} + q^{-1} = 1$ and $A$ is $(\sigma_A(p\theta), \rho_A(p\theta))$-bounded and $S$ is $(\sigma_S(q\theta), \rho_S(q\theta))$-bounded.*

*Proof.* By definition of the dynamic $S$-server we have:

$$\begin{aligned} q(n) &\leq A(n) - \min_{0 \leq k \leq n} \{A(0,k) + S(k,n)\} \\ &= \max_{0 \leq k \leq n} \{A(k,n) - S(k,n)\} \end{aligned}$$

from which we can derive, using Chernoff's inequality[2]:

$$\begin{aligned}\mathbb{P}(q(n) > x) &\leq e^{-\theta x}\mathbb{E}(e^{\theta \max_{0\leq k\leq n}\{A(k,n)-S(k,n)\}}) \\ &\leq e^{-\theta x}\sum_{k=0}^{n}\mathbb{E}(e^{\theta A(k,n)})\mathbb{E}(e^{-\theta S(k,n)}) \\ &\leq e^{-\theta x}e^{\theta(\sigma_A(\theta)+\sigma_S(\theta))}\sum_{k=0}^{n}e^{k\theta(\rho_A(\theta)+\rho_S(\theta))}\end{aligned}$$

Where the independence of $A$ and $S$ has been used in the second line.

Now assume $A$ and $S$ are not stochastically independent and choose $p$ and $q$ like above. We can then replace the independence assumption by using Hölders inequality instead:

$$\mathbb{E}(e^{\theta \max_{0\leq k\leq n}\{A(k,n)-S(k,n)\}}) \leq \sum_{k=0}^{n}\mathbb{E}(e^{\theta A(k,n)}e^{-\theta S(k,n)}) \leq \sum_{k=0}^{n}\mathbb{E}(e^{p\theta(A(k,n)})^{1/p}\mathbb{E}(e^{-q\theta S(k,n)})^{1/q}$$

□

We see in this proof where problems arise, when we encounter uncertainties in the description of the arrival flow $A$. If we do not know about the exact distribution of the increments, we cannot calculate the expression $\mathbb{E}(e^{\theta A(k,n)})$, which in turn prohibits calculation of the backlog bound. Hence, we need the tools of mathematical statistics to bound $\mathbb{E}(e^{\theta A(k,n)})$, which in effect replaces the $(\sigma_A(\theta), \rho_A(\theta))$-bound in the above proof.

## 3. A Framework for a Statistical Network Calculus

In this section, we present the framework of StatNC which operates on past observations of the arrival process, i.e., calculates statistics on the sample $a = (a_{n_0}, \dots, a_{-1})$. Technically, this can simply be seen as a sufficient condition for the employed statistics, which enables calculations of performance bounds while dealing with uncertainties about the arrival process rising from estimations. For brevity denote by $\phi_{m,n}(\theta) := \mathbb{E}(e^{\theta A(m,n)})$ the MGF of $A(m,n)$ at point $\theta$. First, we need a small lemma proving the monotonic behaviour of the backlog bound in the MGF of $A$.

**Lemma 4.** *Let $\hat{\phi}_{m,n}(\theta) \geq \phi_{m,n}(\theta)$ for some $\theta > 0$ and all $m, n \in \mathbb{N}_0$ with $m \leq n$. Assume $A$ being stochastically independent from $S$. Then*

$$\mathbb{P}(q(n) > x) \leq e^{-\theta x}\sum_{k=0}^{n}\hat{\phi}_{k,n}(\theta)\mathbb{E}(e^{-\theta S(k,n)})$$

*for all $n \in \mathbb{N}_0$.*

*If $A$ and $S$ are not stochastically independent we still have:*

$$\mathbb{P}(q(n) > x) \leq e^{-\theta x}\sum_{k=0}^{n}(\hat{\phi}_{k,n}(p\theta))^{1/p}\mathbb{E}(e^{-q\theta S(k,n)})^{1/q}$$

*for some $p$ and $q$ with $p^{-1} + q^{-1} = 1$ and $\hat{\phi}_{k,n}(p\theta) \geq \phi_{m,n}(p\theta)$.*

---

[2]Chernoff's inequality states that for some real random variable $X$ and every $\theta > 0$: $\mathbb{P}(X > x) \leq e^{-\theta x}\mathbb{E}(e^{\theta X})$.

*Proof.* From the proof of theorem 3 we know

$$\mathbb{P}(q(n) > x) \leq e^{-\theta x} \sum_{k=0}^{n} \phi_{k,n}(\theta)\mathbb{E}(e^{-\theta S(k,n)}) \leq e^{-\theta x} \sum_{k=0}^{n} \hat{\phi}_{k,n}(\theta)\mathbb{E}(e^{-\theta S(k,n)})$$

the stochastically dependent case follows in the very same fashion.. □

Define $\mathcal{F}$ to be the space of functions mapping from $\mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{R}^+$ to $\mathbb{R}_0^+$. In expression, if $f \in \mathcal{F}$, then:

$$f : \mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{R}^+ \to \mathbb{R}_0^+$$

An already familiar example for a member of $\mathcal{F}$ is the MGF $\phi_{m,n}(\theta)$ of the arrival flow $A$.

We now provide a theorem how the uncertainties of using statistics can be combined with the probabilistic bounds derived from SNC.

**Theorem 5.** *Let* $\theta^* = \sup\{\theta : \phi_{m,n}(\theta) < \infty\}$ *and* $\Phi : \mathbb{R}^{|n_0|} \to \mathcal{F}$ *be a statistic on* $a = (a_{n_0}, \ldots, a_{-1})$ *such that*

$$\sup_{\theta \in (0,\theta^*)} \mathbb{P}\Big( \bigcup_{m \leq n} \Phi(a)(m,n,\theta) < \phi_{m,n}(\theta) \Big) \leq \alpha.$$

*Then for all* $n \in \mathbb{N}_0$, $\theta < \theta^*$

$$\mathbb{P}(q(n) > x) \leq \alpha + e^{-\theta x} \sum_{k=0}^{n} \Phi(a)(k,n,\theta)\mathbb{E}(e^{-\theta S(k,n)}).$$

*Proof.* Fix some $\theta > 0$:

$$\begin{aligned}
&\mathbb{P}(q(n) > x) \\
=&\mathbb{P}\Big(q(n) > x \cap \bigcup_{m \leq n} \Phi(a)(m,n,\theta) < \phi_{m,n}(\theta)\Big) \\
&+ \mathbb{P}\Big(q(n) > x \cap \bigcap_{m \leq n} \Phi(a)(m,n,\theta) \geq \phi_{m,n}(\theta)\Big) \\
\leq& \alpha + \mathbb{P}\Big(q(n) > x \cap \bigcap_{m \leq n} \Phi(a)(m,n,\theta) \geq \phi_{m,n}(\theta)\Big) \\
\leq& \alpha + \mathbb{P}\Big(q(n) > x \Big| \bigcap_{m \leq n} \Phi(a)(m,n,\theta) \geq \phi_{m,n}(\theta)\Big) \\
\leq& \alpha + e^{-\theta x} \sum_{k=0}^{n} \Phi(a)(m,n,\theta)\mathbb{E}(e^{-\theta S(k,n)})
\end{aligned}$$

□

From the proof, the nature of the condition in the theorem becomes clearer. We need the intersection of the events $\Phi(a)(m,n,\theta) \geq \phi_{m,n}(\theta)$ to leverage from the monotonic behaviour of the backlog bound in terms of the MGF of $A$ (→Lemma 4). We achieve this intersection by partitioning the event $q(n) > x$ and hence have to deal with the corresponding complement, which is the union appearing in the second line of the proof. This union describes the event, that our statistic delivers a value lying below the real MGF of $A$ at least once. We bound this kind of (estimation) error by a confidence level of $\alpha$. The confidence level $\alpha$ can be seen as a parameter of optimization.

Please note, that, for our theorem, we defined the statistic $\Phi$ in the most general way, i.e., being a function on the complete history $a = (a_{n_0}, \dots, a_{-1})$. This does not mean, however, that one has to use all this information to construct a $\Phi$ satisfying the above condition. Assume a subsample $a' = (a'_{n'_0}, \dots, a'_{-1})$ of $a$, such that: $a'_i = a_{j_i}$ for some index $j_i \in \{n_0, \dots, -1\}$ and a statistic $\Phi' : \mathbb{R}^{|n'_0|} \to \mathcal{F}$ on $a'$. If $\Phi'$ meets the assumption of Theorem 5 for $a'$, we can extend it canonically to a statistic $\Phi$, by setting: $\Phi(a) = \Phi'(a')$ for all $a$, such that $a_{j_i} = a'_i$ holds. This shows that using only a part of the history $a$ is just a special case of the above theorem.

Often one may want to give more recent observations a larger impact on the sample and diminish the influence of observations as they get older, e.g., $w_i(a_i) = \beta^{|i+1|}a_i$, $0 < \beta < 1$ ($\rightarrow$exponential smoothing). Such transformations of the sample are also covered by the above theorem: for some weighting function $w : \mathbb{R}^{|n_0|} \to \mathbb{R}^{|n_0|}$ on the sample $a$, the concatenation $\Phi \circ w(a)$ needs to meet the above assumption in order to apply the statistic $\Phi$ on the weighted sample.

Another typical and very practical way of subsampling would be to use a sliding window of length $l$ on the observations, i.e., $a'_i = a_i$, with $i \in \{-l, \dots, -1\}$. The sliding window is particularly interesting for on-line estimation, since it allows the statistical network calculus to dynamically adapt to changes in the arrivals' characteristics. In Section 5, we investigate how the versatility of this dynamic view can be leveraged to achieve better bounds than the static ones of SNC.

For the rest of the report, we stick to the most general notation as in Theorem 5, unless otherwise mentioned.

## 4. Plugging in the Statistical Estimators

As stated above, estimating the quantity $\phi_{m,n}(\theta)$ for an arbitrary $\theta \in (0, \theta^*)$ and $m \leq n \in \mathbb{N}$ is key for StatNC. In the following subsections, we present different scenarios and their corresponding $\phi_{m,n}(\theta)$-estimations as examples how such statistics can be constructed. The crucial point is to meet the condition from Theorem 5. We start with a fairly simple example (exponential i.i.d. increments) to illustrate the idea and then move on to more complex scenarios, which involve non-i.i.d. behaviour of the increments $a_k$ and more advanced statistics.

4.1. **Exponential Traffic.** Assume the $(a_k)_{k \geq n_0}$ to be i.i.d. exponentially distributed for some unknown parameter $\lambda$. The idea to construct $\Phi$ in this scenario, is to estimate $\lambda$ first. For this, note that a lower bound on the real distribution parameter $\lambda$ with confidence level $\alpha$ can be computed by

$$\bar{\lambda} := \frac{\chi^2_\alpha(2|n_0|)}{2 \cdot A(n_0 - 1, -1)},$$

where $\chi^2_\alpha(2|n_0|)$ is the one-sided $\alpha$-quantile of a Chi-Squared distribution with $2|n_0|$ degrees of freedom (scaling $A(n_0 - 1, -1)$ by $2\lambda$ results in a random variable, which is Chi-squared-distributed with $2|n_0|$ degrees of freedom). Moreover, the MGF of the exponential distribution is given by $\phi(\lambda, \theta) = (\frac{\lambda}{\lambda - \theta}) = 1 + (\frac{\theta}{\lambda - \theta})$, for all $\theta < \lambda$. With the simple implication

$$\bar{\lambda} \leq \lambda \Rightarrow \frac{\bar{\lambda}}{\bar{\lambda} - \theta} \geq \frac{\lambda}{\lambda - \theta},$$

we obtain for the case $\bar{\lambda} \leq \lambda$ that

$$\Phi(a)(m,n,\theta) := \left(\frac{\bar{\lambda}}{\bar{\lambda}-\theta}\right)^{n-m} \geq \left(\frac{\lambda}{\lambda-\theta}\right)^{n-m} = \phi_{m,n}(\theta)$$

holds for all $\theta < \lambda$ and $m \leq n \in \mathbb{N}_0$.

Hence we obtain

$$\begin{aligned} 1-\alpha &= \mathbb{P}(\bar{\lambda} \leq \lambda) \\ &\leq \inf_{\theta \in (0,\lambda)} \mathbb{P}\Big( \bigcap_{m \leq n} \Phi(a)(m,n,\theta) \geq \phi_{m,n}(\theta) \Big). \end{aligned}$$

Or, equivalently

$$\sup_{\theta \in (0,\lambda)} \mathbb{P}\Big( \bigcup_{m \leq n} \Phi(a)(m,n,\theta) < \phi_{m,n}(\theta) \Big) \leq \alpha,$$

which is exactly the condition of Theorem 5.

4.2. **Bandwidth-Limited i.i.d. Traffic .** Assume again the increments of $A$ to be i.i.d., now also adhering to a bandwidth limitation $M$, i.e., no more than $M$ data units per time slot can arrive. The latter is a valid assumption as any real access link has such a restriction. In contrast to the previous subesection, we do not have any further knowledge about the distribution of the increments $a_k$. Yet, it turns out that we still have enough information to construct a statistic $\Phi$ fitting Theorem 5. The Dvoretzky-Kiefer-Wolfowitz inequality [22] is very useful in this setup:

**Lemma 6.** *Let* $F_{n_0}(x) := {}^1\!/_{|n_0|} \sum_{k=n_0}^{-1} \mathbf{1}_{\{a_k \leq x\}}$ *be the empirical distribution function of the sample* $a = (a_{n_0}, \ldots, a_{-1})$ *and* $F$ *be the distribution of* $A$ *. Then we have for all* $\varepsilon > 0$ *the following:*

$$\mathbb{P}\Big( \sup_{x \in [0,M]} |F_{n_0}(x) - F(x)| \leq \varepsilon \Big) \geq 1 - 2e^{-2|n_0|\varepsilon^2}.$$

Since the arrivals are bounded we can fix some arbitrary $\theta > 0$ for the rest of this subsection. The next theorem constructs $\Phi$:

**Theorem 7.** *Let* $\varepsilon > 0$*. The statistic* $\Phi$ *defined by*

$$\Phi(a)(m,n,\theta) := (\bar{A} + \varepsilon(e^{\theta M} - 1))^{n-m}$$

*satisfies the condition in Theorem 5. Here*

$$\bar{A} := \frac{1}{|n_0|} \sum_{k=n_0}^{-1} e^{\theta a_k}.$$

*Proof.* From the event in the Dvoretzky-Kiefer-Wolfowitz inequality (→ Lemma 6), we can derive successively:

$$F(x) \geq F_{n_0}(x) - \varepsilon \qquad \forall\, x \in [0,M],$$

$$1 - F(x) \leq 1 - F_{n_0}(x) + \varepsilon \qquad \forall\, x \in [0,M],$$

$$\mathbb{P}\left(e^{\theta A} > x\right) \leq 1 - F_{n_0}\left({}^1\!/_\theta \log(x)\right) + \varepsilon \qquad \forall\, x \in [1, e^{\theta M}],$$

$$\begin{aligned}\mathbb{E}\left(e^{\theta A}\right) &= 1 + \int_1^{e^{\theta M}} \mathbb{P}\left(e^{\theta A} > x\right)\, dx \\ &\leq 1 + \int_1^{e^{\theta M}} 1 - F_{n_0}\left({}^{1}\!/\!_{\theta} \log(x)\right) + \varepsilon\, dx.\end{aligned}$$

Hence, we have for all $\theta > 0$:

$$\begin{aligned}&\mathbb{P}\Big(\mathbb{E}\left(e^{\theta A}\right) \leq 1 + \int_1^{e^{\theta M}} 1 - F_{n_0}\Big(\tfrac{1}{n}\log(x)\Big) + \varepsilon\, dx\Big) \\ \geq\, &\mathbb{P}\Big(\sup_{x\in[0,M]} |F_{n_0}(x) - F(x)| \leq \varepsilon\Big) \\ \geq\, &1 - 2e^{-2|n_0|\varepsilon^2}.\end{aligned}$$

This means we have constructed a one-sided confidence interval for $\mathbb{E}(e^{\theta A})$ with significance level $\alpha = 2e^{-2|n_0|\varepsilon^2}$, which is given by:

$$\left[0, 1 + \int_1^{e^{\theta M}} 1 - F_{n_0}\left({}^{1}\!/\!_{\theta}\log(x)\right) + \varepsilon\, dx\right].$$

The appearing integral can be simplified by:

$$\begin{aligned}\int_1^{e^{\theta M}} 1 - F_{n_0}\left({}^{1}\!/\!_{\theta}\log(x)\right) + \varepsilon\, dx &= (1+\varepsilon)(e^{\theta M} - 1) - {}^{1}\!/\!_{|n_0|} \int_1^{e^{\theta M}} \sum_{i=n_0}^{-1} \mathbf{1}_{\{e^{\theta a_i} \leq x\}} dx \\ &= (1+\varepsilon)(e^{\theta M} - 1) - {}^{1}\!/\!_{|n_0|}(e^{\theta M} - e^{\theta a_{(|n_0|)}} + \ldots + e^{\theta M} - e^{\theta a_{(1)}}) \\ &= \bar{A} - 1 + \varepsilon(e^{\theta M} - 1)\end{aligned}$$

where $a_{(i)}$ is the $i$-th order statistic of the sample and $\bar{A}$ the sample mean of the $e^{\theta a_i}$

Inserting the corresponding $\varepsilon$ for a significance level $\alpha$, the confidence interval becomes:

$$\left[0, \bar{A} + \sqrt{\frac{-\log(\frac{\alpha}{2})}{2|n_0|}}(e^{\theta M} - 1)\right].$$

For the statistic $\Phi$ we indeed have:

$$\begin{aligned}&\inf_\theta \mathbb{P}\Big(\bigcap_{m,n} \Phi(a)(m,n,\theta) \geq \phi_{m,n}(\theta)\Big) \\ =\, &\inf_\theta \mathbb{P}\Big(\bigcap_{m,n} \prod_{k=m+1}^{n} \bar{A} + \varepsilon(e^{\theta M} - 1) \geq \prod_{k=m+1}^{n} \mathbb{E}(A)\Big) \\ \geq\, &\inf_\theta \mathbb{P}(\bar{A} + \varepsilon(e^{\theta M} - 1) \geq \mathbb{E}(A)) \\ \geq\, &1 - \alpha.\end{aligned}$$

Or equivalently:

$$\sup_\theta \mathbb{P}\Big(\bigcup_{m,n} \Phi(a)(m,n,\theta) < \phi_{m,n}(\theta)\Big) \leq \alpha.$$

Hence, again, now for bandwidth-limited i.i.d. arrivals, the statistic $\Phi(a)(m,n,\theta) := (\bar{A} + \varepsilon(e^{\theta M} - 1))^{n-m}$ satisfies the condition of Theorem 5 and can thus be used to calculate the desired performance bounds. □

4.3. **Markov-Modulated Arrivals.** Next, we discuss a traffic class, in which the i.i.d. assumption is dropped and for which multiple statistics are combined to construct $\Phi$. For this consider a Markov-modulated arrival, with a Markov chain $(Y_k)_{k\in\{n_0,\dots\}}$ corresponding to the transition matrix

$$T = \begin{pmatrix} \mu & 1-\mu \\ 1-\nu & \nu \end{pmatrix}$$

and denote the first (second) state as *On*-state (*Off*-state). The increments of the arrival $A$ are now defined by:

$$a_k = \begin{cases} 0 & \text{if } Y_k \text{ is in } \textit{Off}\text{-state} \\ x_k & \text{if } Y_k \text{ is in } \textit{On}\text{-state} \end{cases}$$

where $(x_k)_{k\in\{n_0,\dots\}}$ is a sequence of i.i.d. random variables bounded by the bandwidth limitation $M$. Note that the increments $a_k$ are neither identically distributed nor stochastically independent. This model generalizes the well-known and popular Markov-modulated On-Off traffic model [1], with the difference, that in the *On*-state the arrivals are defined by a random process, instead of a constant rate.

Before we can construct $\Phi$, we need a lemma, showing the monotonicty of $\phi_{m,n}(\theta)$ with respect to the parameters $\mu$ and $\nu$. For this define $\theta^* = \sup\{\theta : \mathbb{E}(e^{\theta x_k}) < \infty\}$.

**Lemma 8.** *For the above model, all $m \le n \in \mathbb{N}$ and all $\theta \in (0, \theta^*)$ it holds that*

$$\mu \ge \tilde{\mu} \quad \Rightarrow \quad \mathbb{E}(e^{\theta A_{\mu,\nu}(m,n)}) \le \mathbb{E}(e^{\theta A_{\tilde{\mu},\nu}(m,n)}),$$

*and*

$$\nu \le \tilde{\nu} \quad \Rightarrow \quad \mathbb{E}(e^{\theta A_{\mu,\nu}(m,n)}) \le \mathbb{E}(e^{\theta A_{\mu,\tilde{\nu}}(m,n)}).$$

*Proof.* We start with the first statement. Without loss of generality we assume that the chain starts in the *Off*-state. We fix some $m$, $n \in \mathbb{N}$ such that $m \le n$ and use the notation $A_{\tilde{\mu},\nu}(m,n) =: \tilde{A}(m,n)$. First note, that it suffices to show that $\mathbb{P}\left(O(n-m) \ge k\right) \le \mathbb{P}\left(\tilde{O}(n-m) \ge k\right)$ holds for all $k = 1, \dots, n-m$, where $O(m,n)$ denotes the number of times the signal $Y$ is in the *On*-state during times $m+1, \dots, n$. ($\tilde{O}(m,n)$ is connected corespondingly to the signal of the altered chain.) Indeed we would have for some $x \in \mathbb{R}^+$ :

$$
\begin{aligned}
&\mathbb{P}\left(e^{\theta A(m,n)} \geq x\right)\\
=&\sum_{k=0}^{n-m} \mathbb{P}\left(O(m,n)=k\right)\mathbb{P}\left(e^{\theta A(m,n)} \geq x \,\middle|\, O(m,n)=k\right)\\
=&\sum_{k=0}^{n-m} \mathbb{P}\left(O(m,n)=k\right)\mathbb{P}\left(e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(k)})} \geq x\right)\\
=&\sum_{k=0}^{n-m} \mathbb{P}\left(O(m,n)=k\right)\left(\mathbb{P}\left(e^{\theta\cdot 0}\geq x\right)+\sum_{l=1}^{k}\mathbb{P}\left(e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l)})} \geq x \cap e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l-1)})} < x\right)\right)\\
=&\sum_{k=0}^{n-m} \mathbb{P}\left(O(m,n)=k\right)\mathbb{P}\left(e^{\theta\cdot 0}\geq x\right)\\
&\quad+\sum_{l=1}^{n-m}\mathbb{P}\left(e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l)})} \geq x \cap e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l-1)})} < x\right)\sum_{k=l}^{n-m}\mathbb{P}\left(O(m,n)=k\right)\\
=&\mathbb{P}\left(O(m,n)\geq 0\right)\mathbb{P}\left(e^{\theta\cdot 0}\geq x\right)\\
&\quad+\sum_{l=1}^{n-m}\mathbb{P}\left(e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l)})} \geq x \cap e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l-1)})} < x\right)\mathbb{P}\left(O(m,n)\geq l\right)\\
\leq&\mathbb{P}\left(\tilde{O}(m,n)\geq 0\right)\mathbb{P}\left(e^{\theta\cdot 0}\geq x\right)\\
&\quad+\sum_{l=1}^{n-m}\mathbb{P}\left(e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l)})} \geq x \cap e^{\theta(a_{On}^{(1)}+\ldots+a_{On}^{(l-1)})} < x\right)\mathbb{P}\left(\tilde{O}(m,n)\geq l\right)\\
=&\mathbb{P}\left(e^{\theta\tilde{A}(m,n)}\geq x\right)
\end{aligned}
$$

and by the monotonicity of the integral:

$$
\mathbb{E}\left(e^{\theta A(m,n)}\right)=\int_{\mathbb{R}}\mathbb{P}\left(e^{\theta A(m,n)}\geq x\right)dx \leq \int_{\mathbb{R}}\mathbb{P}\left(e^{\theta\tilde{A}(m,n)}\geq x\right)dx=\mathbb{E}\left(e^{\theta\tilde{A}(m,n)}\right)
$$

To show $\mathbb{P}\left(O(n-m)\geq k\right)\leq\mathbb{P}\left(\tilde{O}(n-m)\geq k\right)$ for all $k=0,\ldots,n-m$, realize the Markov chain in the following way: In each time slot $k$ we sample random variable $U_k$, being i.i.d. uniformly distributed on $[0,1]$ and define the signal $Y$ by:

$$
Y_{k+1}=\begin{cases} On & \text{if } U_k\in(\mu,1] \text{ and } Y_k=Off\\ Off & \text{if } U_k\in[0,\mu] \text{ and } Y_k=Off\\ On & \text{if } U_k\in[0,\nu] \text{ and } Y_k=On\\ Off & \text{if } U_k\in(\nu,1] \text{ and } Y_k=On\end{cases}
$$

(the signal $\tilde{Y}$ is realized in the same way, with $\mu$ replaced by $\tilde{\mu}$)

Denote now by $\mathcal{P}_k$ the set of all sequences $[s_l]_{l=0,\ldots,n-m}\in\{On,\,Off\}^{n-m}$ for which holds

$$
\sum_{l=1}^{n-m} I_{\{s_l=On\}}=k
$$

In words this just means, that if a trajectory of the signal space lies in $\mathcal{P}_k$ it visits the *On*-state exactly $k$ times up to timestep $n-m$. With the help of the above method of realizing the Markov chain, we can give for every trajectory $P \in \mathcal{P}_k$ a cylinder-set $A_P \subset [0,1]^{n-m}$, such that

$$\mathbb{P}(\text{the signal follows } P) = \mathbb{P}\left(\prod_{l=0}^{n-m-1} U_l \in A_P\right)$$

It is important to note here, that the probability $\mathbb{P}\left(\prod_{l=0}^{n-m-1} U_l \in A_P\right)$ is invariant with respect to permutations of the order of the random variables $U_l$. I.e. for every bijection $\sigma : \{0,\dots,n-m\} \to \{0,\dots,n-m\}$ holds

$$\mathbb{P}\left(\prod_{l=0}^{n-m-1} U_l \in A_P\right) = \mathbb{P}\left(\prod_{l=0}^{n-m-1} U_{\sigma(l)} \in A_P\right).$$

We will see, that for a fitting permutations we can conclude $\mathbb{P}\left(\prod_{l=0}^{n-m-1} U_{\sigma(l)} \in A_P\right) \leq \mathbb{P}\left(\tilde{O}(m,n) \geq k\right)$. To find this permutation define the following time-indices:

$$\begin{aligned} u_1 &= \min\{l \,|\, U_l \in (\mu, 1]\} \\ \tilde{u}_1 &= \min l \,|\, U_l \in (\tilde{\mu}, 1] \end{aligned}$$

Define as long as $d_i$, $u_i$, $\tilde{u}_i \leq n-m-1$ recursively

$$\begin{aligned} d_i &= \min\{l > u_i \,|\, U_l \in (\nu, 1]\} \\ u_i &= \min\{l > d_{i-1} \,|\, U_l \in (\mu, 1]\} \\ \tilde{u}_i &= \min\{l > d_{i-1} \,|\, U_l \in (\tilde{\mu}, 1]\} \end{aligned}$$

and denote by $I$ the largest subindex $i$ for which $u_i \leq n-m-1$. Note that $\tilde{u}_i \leq u_i$, by the above structure of the transition matrix and the assumption that $\mu \geq \tilde{\mu}$. Define a sequence of permutations $[\sigma_i]_{i \leq I}$:

$$\begin{aligned} \sigma_i(l) &= l && \text{if } l \leq \tilde{u}_i \\ \sigma_i(l) &= l - (u_i - \tilde{u}_i) && \text{if } l > u_i \\ \sigma_i(l) &= n-m-1+l-u_i && \text{if } \tilde{u}_i < l \leq u_i \end{aligned}$$

This permutation does the following with the path: It checks if the modified Markov chain would jump earlier to the *On*-state, than the original Markov chain. If it does, the time slots, which lie between the jump of the modified and the original chain, are moved to the end of the timeline, while all later time slots move up to fill the just produced gap. By this the number of timeslots corresponding to *On*-states can not decrease (only "*Off*-states" are moved).

Denote by $\sigma$ now the permutation, which results by successively applying the above permutations $\sigma = \sigma_I \circ \sigma_{I-1} \ldots \circ \sigma_1$. For each $\omega \in \{\omega \mid \prod_{l=0}^{n-m-1} U_{\sigma(l)} \in A_P\}$ holds for the modified transition matrix

$$\tilde{O}(m,n)(\omega) \geq k$$

since the modified trajectory has at least as many visits in *On*-states than the original walk in the unpermutated trajectory. Hence it holds for every path $P \in \mathcal{P}_k$:

$$\mathbb{P}\left(\prod_{l=0}^{n-m-1} U_l \in A_P\right) = \mathbb{P}\left(\prod_{l=0}^{n-m-1} U_{\sigma(l)} \in A_P\right) \leq \mathbb{P}\left(\tilde{O}(m,n) \geq k\right)$$

This allows us to finish the prove for the first statement, by using:

$$\begin{aligned}\{O(m,n) \geq k\} &= \bigcup_{l=k}^{n-m} \{O(m,n) = l\} \\ &\subset \bigcup_{l=k}^{n-m} \{\tilde{O}(m,n) \geq l\} \\ &= \{\tilde{O}(m,n) \geq k\}\end{aligned}$$

The second statement follows very similar. Define $P(m,n)$ and $\tilde{P}(m,n)$ as the number of times the Markov chain and the modified Markov chain are in the *Off*-state during $(m,n]$. It can be shown very similar to the previous that:

$$\mathbb{P}\left(\tilde{P}(m,n) > k\right) \leq \mathbb{P}\left(P(m,n) > k\right) \qquad \forall\, k \in \{0,\dots,n-m\}$$

From this follows

$$\begin{aligned}\mathbb{P}\left(n-m-\tilde{P}(m,n) < n-m-k\right) &\leq \mathbb{P}\left(n-m-P(m,n) < n-m-k\right) \\ \mathbb{P}\left(\tilde{O}(m,n) < n-m-k\right) &\leq \mathbb{P}\left(O(m,n) < n-m-k\right) \\ \mathbb{P}\left(\tilde{O}(m,n) \geq n-m-k\right) &\geq \mathbb{P}\left(O(m,n) \geq n-m-k\right)\end{aligned}$$

for all $k \in \{0,\dots,n-m\}$. From which we can follow $\mathbb{E}\left(e^{\theta A_{\mu,\nu}(m,n)}\right) \leq \mathbb{E}\left(e^{\theta A_{\mu,\tilde{\nu}}(m,n)}\right)$ as in the proof of the first statement. □

We next construct a $\Phi$ in this case of Markov-modulated arrivals, for which neither the transition probabilities $\mu$ and $\nu$, nor the distribution of the $(x_i)$ are known. Of course, we want this statistic to satisfy the requirements of our framework again, i.e., the condition in Theorem 5. For an arrival sample $a = (a_{n_0},\dots,a_{-1})$, let $X_{0,0}, X_{0,1}, X_{1,0}$, and $X_{1,1}$ denote the observed number of transitions from the *Off*-state to *Off*-state, from *Off*-state to *On*-state, from *On*-state to *Off*-state and from *On*-state to *On*-state, respectively. Further denote the observed number of visits in the *On*-state by $O(n_0,n_{-1}) = O$ and the number of visits in the *Off*-state by $P(n_0,n_{-1}) = P$.

**Theorem 9.** *Choose some confidence level $\alpha = \alpha_\mu + \alpha_\nu + \alpha_d$ and consider some sample $a = (a_{n_0},\dots,a_{-1})$ with $O \neq 0$. Define the statistics*

$$\begin{aligned}\mu_l &:= \beta^{-1}(\alpha_\mu; X_{0,0}, P - X_{0,0} + 1) \\ \nu_u &:= \beta^{-1}(1-\alpha_\nu; X_{1,1}+1, O - X_{1,1})\end{aligned}$$

*where $\beta^{-1}$ is the inverse of the beta-distribution. Further define the transition matrix*

$$T^* = \begin{pmatrix} \mu_l & 1-\mu_l \\ 1-\nu_u & \nu_u \end{pmatrix}.$$

*Define*

$$A^* = \bar{A} + \left(\frac{-\log\left({}^{\alpha_d}\!/_2\right)}{2|O|}\right)^{1/2}\left(e^{\theta M} - 1\right).$$

*Then the statistic* $\Phi : \mathbb{R}^{|n_0|} \to \mathcal{F}$ *defined by:*

$$\Phi(a)(m,n,\theta) := A^* \frac{\bar{x}_{On} \vee \bar{x}_{Off}}{\bar{x}_{On} \wedge \bar{x}_{Off}} \rho(\bar{E}T^*)^{n-m-1}$$

*satisfies the condition of the statistical framework theorem. Here* $\bar{A} := \frac{1}{O}\sum_{i:Y_i=On} e^{\theta a_i}$, $\bar{E}$, $\bar{x}_{On}$ *and* $\bar{x}_{Off}$ *are from the* $\sigma(\theta), \rho(\theta)$*-bound for Markov-modulated arrivals (*$\to$ *Appendix 6) with fixed arrivals in the On-states equal to* $1/\theta \log(A^*)$*.*

*Proof.* First we show that $[\mu_l, 1]$ and $[0, \nu_u]$ are confidence intervals for $\mu$ and $\nu$ and the confidence levels $\alpha_\mu$ and $\alpha_\nu$, respectively. Basically, these are the well-known Clopper-Pearson intervals [9], which are constructed as follows: Interpret $X_{0,0}$ as the number of successes in a $Bin(\mu, P)$-distributed random variable. It is known that for some $X \sim Bin(p, n)$ it holds that

$$\mathbb{P}(X \leq k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} = \beta(1-p;\, n-k,\, k+1).$$

From that we can continue with

$$\beta(1-p,\, n-k,\, k+1) = 1 - \beta(p;\, k+1,\, n-k).$$

We now ask for the smallest $\mu'$ such that a random variable $X \sim Bin(\mu', P)$ meets:

$$\mathbb{P}_{\mu'}(X \geq X_{0,0}) \geq \alpha_\mu,$$

or, equivalently

$$\mathbb{P}_{\mu'}(X \leq X_{0,0} - 1) \leq 1 - \alpha_\mu.$$

To find $\mu'$ consider

$$1 - \beta(\mu';\, X_{0,0},\, P - X_{0,0} + 1) = \mathbb{P}_{\mu'}(X \leq X_{0,0} - 1) \overset{!}{=} 1 - \alpha_\mu,$$

which is solved by $\mu_l = \beta^{-1}(\alpha;\, X_{0,0},\, P - X_{0,0} + 1)$. Using a simple coupling argument and the definition of $\mu_l$ one obtains the implication:

$$\mathbb{P}(\mu_l > \mu) \leq \mathbb{P}(X \geq X_{0,0}) = \alpha_\mu.$$

Very similarly we obtain:

$$\mathbb{P}(\nu_u < \nu) \leq \mathbb{P}(X \leq X_{1,1}) = \alpha_\nu,$$

where $X$ is now a $Bin(\nu_u, O)$-distributed variable.

Now fix some arbitrary $\theta \in (0, \theta^*)$ and assume for the moment

$$\mu_l < \mu,$$
$$\nu_u > \nu,$$
$$A^* < \mathbb{E}(e^{\theta a_{On}}).$$

Then for all $m, n \in \mathbb{N}$ it would hold that:

$$\mathbb{E}(e^{\theta A(m,n)}) \leq \mathbb{E}(e^{\theta A_{\mu_l,\nu}(m,n)}) \leq \mathbb{E}(e^{\theta A_{\mu_l,\nu_u}(m,n)}) \leq \Phi(a)(m,n,\theta),$$

where we still have to show the last inequality. Putting the proof of the last inequality on hold, we henceforth have for all $\theta \in (0, \theta^*)$

$$\begin{aligned}\alpha = \alpha_\mu + \alpha_\nu + \alpha_d &\geq \mathbb{P}(\mu_l > \mu \cup \nu_u < \nu \cup A^* < \mathbb{E}(e^{\theta a_{On}})) \\ &\geq \mathbb{P}\Big(\bigcup_{m \leq n} \mathbb{E}(e^{\theta A(m,n)}) > \Phi(a)(m,n,\theta)\Big)\end{aligned}$$

which is what we wanted to show.

To show the missing inequality define a new Markov-modulated arrival with a constant rate $a^*_{On} = {}^1\!/\!_\theta \log(A^*)$ and $T^*$ as transition matrix. We then have $\mathbb{E}(e^{\theta a^*_{On}}) = A^*$ and the $\sigma(\theta), \rho(\theta)$-bound is given by $\Phi(a)(m,n,\theta)$. Further we have:

$$\begin{aligned}\mathbb{E}(e^{\theta A(m,n)}) &= \sum_{k=0}^{n-m} \mathbb{P}(O(m,n) = k)\mathbb{E}(e^{\theta a_{On}})^k \\ &\leq \sum_{k=0}^{n-m} \mathbb{P}(O(m,n) = k)\mathbb{E}(e^{\theta a^*_{On}}) \\ &\leq \Phi(a)(m,n,\theta).\end{aligned}$$

Here, Lemma 8 was used in the second line. This completes the proof. □

4.4. **Summary.** We provided three examples with differing degrees of assumed knowledge and complexity when constructing the statistic $\Phi$. From the formulation of the framework it is clear, that the technically hard part in applying the StatNC lies in constructing such estimators. Taking care of other, potentially more complex arrival processes is hence just a question of finding the corresponding $\Phi$, i.e., meeting the condition of the framework theorem. Admittedly, this can be hard in some circumstances and, for example, we leave the construction of $\Phi$ for long-range dependent traffic types for future work.

## 5. Numerical Evaluation – StatNC at Work

In this section, we compare the statistical network calculus to its stochastic counterpart. For this we investigate how high the costs of involving statistics are (in terms of looser bounds). Furthermore, we study special properties of StatNC which the SNC lacks; these are its dynamic view on the measurements, as well as its robustness against false assumptions.

**Scenario 1: The Price of StatNC.** In our first scenario, we study if the additional uncertainties resulting from the statistical part of the performance bounds are acceptable. In expression, we calculate the smallest $b$ such that we can still guarantee

$$\mathbb{P}(q(n) > b) \leq \varepsilon,$$

with our methods of StatNC (or with the methods of standard SNC). For a perfect bound, we would encounter after a large number $N$ of simulations that roughly $N \cdot \varepsilon$ of the simulations produce a backlog larger than $b$ at time $n$. Hence, we simulate the backlog process for time $n$ in $N$ repetitions and compare the empirical distribution of the observed backlogs with the $b$ we found from the above formula above. The bounds are the better, the closer they lie to the $(1-\varepsilon)$-quantile of the empirical backlog distribution.

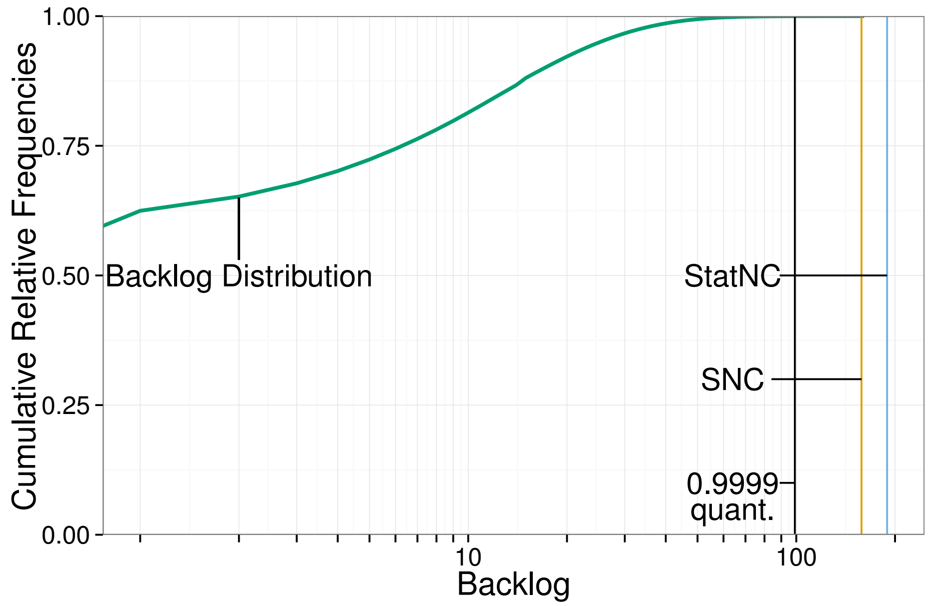


Figure 5.1. StatNC and SNC backlog bounds as well as empirical backlog distribution of the backlog measured at time $n = 100$ for $N = 10^6$ simulation runs.

To that end, we simulate a Markov-modulated arrival process, as described in Section 4.3, with $x_i$ being exponentially distributed, but capped by a bandwidth limitation $M$. The parameter $\lambda$ of the exponential distributions is chosen to be 0.2, while the bandwidth limitation is set to $M = 20$ (which means a hypothetical access link is maximally utilized at 25%). The transition probabilities of the Markov chain are given by $\mu = 0.9$ and $\nu = 0.9$. We use a constant rate server with rate $c = 5$, which means during the *On*-state, considering the bandwidth limitation, we see a peak utilization of $\frac{1}{\lambda c}(1 - e^{-\lambda M}) \approx 98\%$. Considering the bandwidth limitation and the fact, that we are not always in the *On*-state, we compute an average utilization of roughly $\approx$ 49%. The computation of the backlog bound based on SNC follows Appendix [REF], while the StatNC bounds are computed according to Section 4.3. For illustration, we have simulated $N = 10^6$ runs of this system and evaluated the backlog at time slot 100 (at which time in all simulation runs the initial distribution of the Markov chain had faded out and steady-state was reached). In Figure 5.1, the empirical distribution function of the backlog is plotted; for the bounds a violation probability of $\varepsilon = 10^{-4}$ was assumed. As can be observed, both bounds are reasonably close to the $(1 - \varepsilon)$-quantile, but, even more importantly, the bounds are pretty close to each other. This demonstrates that the price we pay for using StatNC is not too high.

**Scenario 2: Exploiting the Dynamic Behaviour of StatNC.** In the second scenario, we investigate StatNC's dynamic point of view. In particular, we use a sliding window approach over the last $l$ observations (as discussed in Section 3). Using this kind of sub-sampling, we eventually forget old measurements and "learn" from new arrivals instead. As such, the observation window allows to track changes in the arrival process (stemming, e.g., from non-stationarities such as time-of-day or other seasonal effects), which take place on longer timescales. For example, if we imagine a flow starting with large increments and diminishing over time, standard SNC faces problems; it lacks the adaptability to track this behaviour and its bounds get looser over time. On the other hand, StatNC can adapt by forgetting about the first large increments as time passes.

To investigate this effect, we use a Markov-modulated arrival process, similar to the previous one, but instead of having *On*- and *Off*-states, we use states *High* and *Low*. For both of these, arrivals are drawn from an exponential distribution with a parameter $\lambda_{x_i}$ (and then capped by $M$); here, the parameter $\lambda_{x_i}$ depends on the state of the Markov chain (*High* or *Low*).

In this scenario, we use the estimators from Subsections 4.1 and 4.2 and not the estimator presented in Subsection 4.3. The goal is, instead of "learning" the Markov chain itself, to use the observation window for tracking changes of states. Therefore, we use transition probabilities of $\mu = 0.999$ and $\nu = 0.999$; this means in expectation we stay 1000 time slots in one of the states until we change into the other one; this emulates a non-stationary behaviour of the arrival process. Further, we set $\lambda_{Low} = 5$ and $\lambda_{High} = 0.2$ and a bandwidth limitation of $M = 10$. With a service rate of $c = 5$, we have–taking the bandwidth limitation into account–a utilization of 4%, while residing in the *Low*-state and 86% in the *High*-state. In the simulations, we started the arrival process at time $n_0 = -1000$ to provide StatNC with an initial observation window. A typical run of this scenario is plotted in Figures 5.2 (for the exponential traffic estimator) and 5.3 (for the i.i.d. bandwidth-limited estimator). In addition to the bounds, we have plotted the simulated backlog process over time, to see how close the bounds are. Due to their dynamic nature, the StatNC bounds also evolve over time. Like the SNC bounds, they are computed for a violation probability of $\varepsilon = 10^{-4}$ and for a time which lies $n$ time slots after the point they have been computed; we provide the results for $n = 10$ and $n = 1000$ time slots, representing a short and long prediction horizon, respectively. As can be seen clearly, the bounds react and ultimately adapt to the observed arrivals: If the arrival intensity is high (indicated by larger backlogs), the statistical bounds also increase, while they decrease, when the Markov chain changes to the *Low*-state. One can also observe that the StatNC bounds track the changes of states with some delay, since old measurements need to be discarded from the observation window first. The effect of the prediction horizon $n$ is such that larger $n$ result in higher bounds irrespective of the bounding method. For comparison, we also provide the SNC bounds calculated by exactly modelling the Markov chain (→Appendix 6). As can be observed, although the SNC uses complete information its bounds lie far above the StatNC bounds, which perform very well in terms of staying close, but not being violated too often (in accordance with the violation probability $\varepsilon = 10^{-4}$).

Another effect that can be observed, when comparing the two plots with each other, is how helpful the additional information about the exponential distribution is. We see, for the same run, that the StatNC bounds using knowledge about the type of distribution in the *High*-state, performs moderately better, compared to the non-parametric i.i.d. estimator of Subsection 4.2. We will see in the next scenario, however, that taking more assumptions about the arrivals into account, bears the risk of making false assumptions, which in turn can be fatal for the bounds.

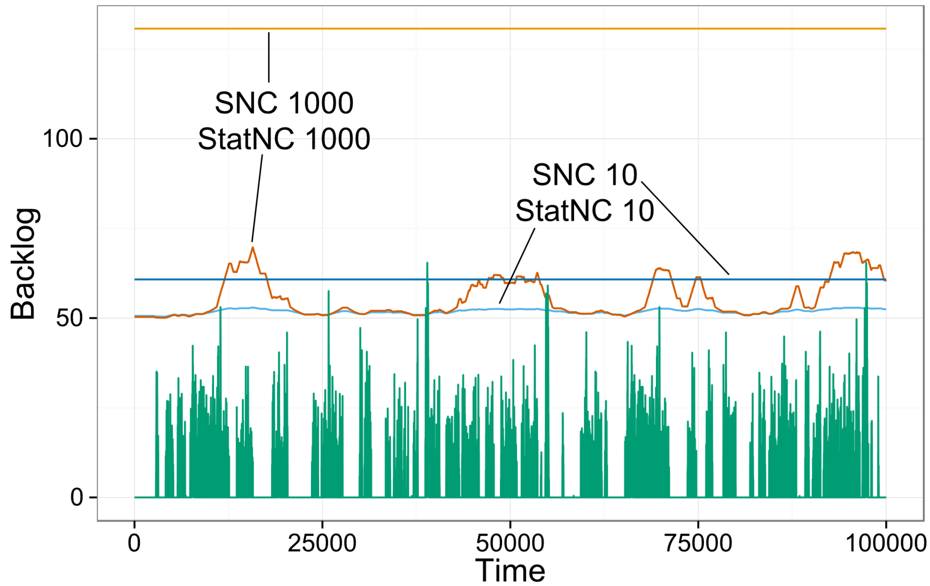


FIGURE 5.2. The backlog process for a typical simulation run as well as different StatNC and SNC bounds for $n = 10, 1000$. Here, the parametric estimator of Subsection 4.1 was used for the StatNC bounds.

**Scenario 3: Robustness of StatNC.** In the third scenario, we investigate the robustness of StatNC and SNC bounds against false assumptions on the arrival process. This reveals another feature of StatNC when using the estimator of Subsection 4.2: StatNC can cope with rather few assumptions about the arrivals and is therefore more robust than SNC.

To illustrate this, we let SNC make a false assumption about the distribution of the i.i.d. increments of the arrival process: For this we have chosen the increments to be i.i.d. Pareto distributed with parameters $x_{min}$ and $s$, again capped by the bandwidth limitation $M$; yet, for the calculation of the SNC bound we assume the increments to be i.i.d. exponentially distributed with parameter $\lambda$ (again capped by $M$), with $\lambda$ set such that the expectations of the Pareto-distributed arrivals and the assumed exponentially distributed arrivals coincide (for details see Appendix [REFIII]). On the other hand, the StatNC using the non-parametric estimator from Subsection 4.2 cannot–by definition–make such a false assumption about the arrivals.

In Figure 5.4, it can be observed that false assumptions on the arrival process can lead to desastrous results. As in the first scenario, the empirical distribution of the backlog is displayed and compared against the bounds calculated by StatNC and SNC. In the plot, parameters have been set as $x_{min} = 1$, $s = 1$, $M = 55$; further, we used a violation probability of $\varepsilon = 10^{-4}$ at time $n = 100$. The plot shows the empirical backlog distribution for $10^6$ simulation runs, which in turn means, we would expect a tight $10^{-4}$ bound to be violated 100 times in expectation. The SNC bound however is broken by 234,526 runs, i.e., in approximately 23% of the

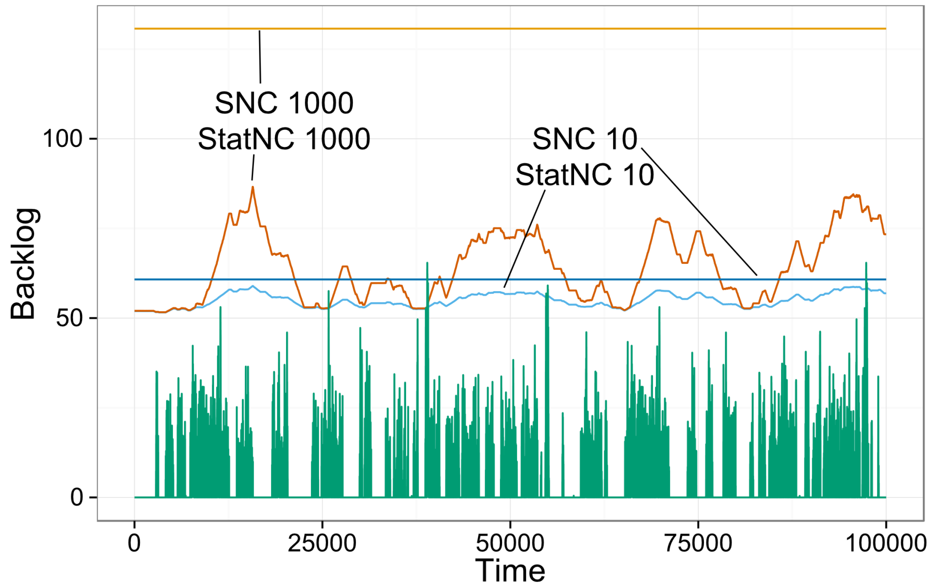


Figure 5.3. The backlog process for a typical simulation run as well as different StatNC and SNC bounds for $n = 10$, 1000. Here, the non-parametric estimator of Subsection 4.2 was used for the StatNC bounds.

simulations! This lies far below the empirical $(1-\varepsilon)$-quantile, the location of a sharp bound. This means the SNC is far too optimistic and hence is rendered useless. In contrast, the StatNC remains valid and stays reasonably close to the empirical quantile–a very satisfying result.

## 6. Conclusion

By integrating statistical methods into the network calculus framework in order to deal with the frequent uncertainty about arrivals, we believe to have made an important step towards a better applicability of network calculus. In particular, the dynamic mode of operation of the newly developed statistical network calculus is attractive for many application fields where uncertainty and permanent change rules and modelling assumptions are already outdated when the actual system is under operation. After providing the basic technical results for StatNC–a framework theorem providing a sufficient condition for statistical estimators of the arrival process to connect them with the SNC framework and several matching estimators–we were able to make a case for the promising opportunities of the novel StatNC framework in a set of numerical experiments.

Given the positive results from this report, there are many opportunities for future work within the StatNC framework: besides the already mentioned estimator for long-range dependent traffic, there are many more useful estimators that can be conceived; also, more sophisticated sub-sampling techniques than sliding windows, e.g., optimally weighted estimators, could provide even better reaction times; and

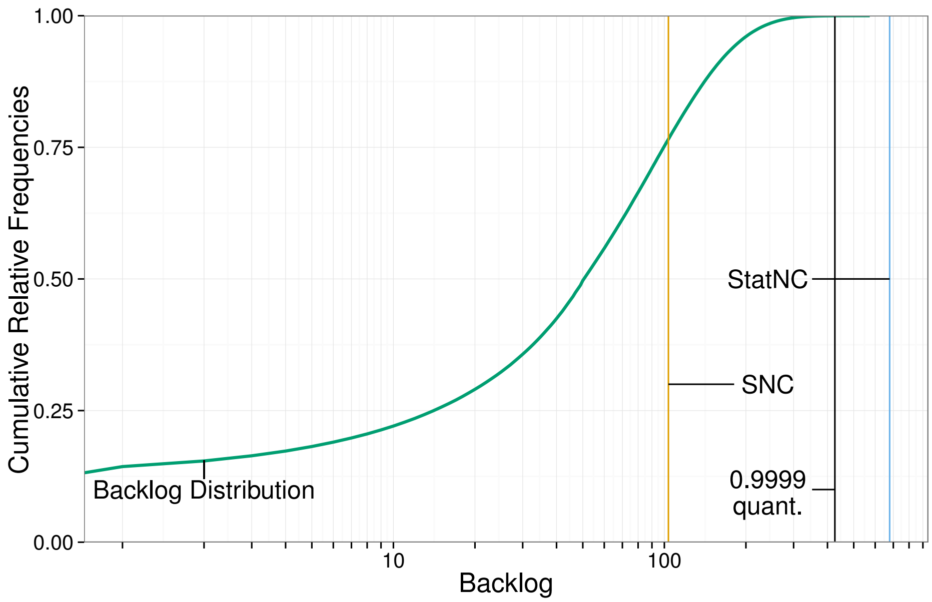


FIGURE 5.4. Under a false distribution assumption (exponential instead of Pareto distributed increments), SNC delivers a grossly invalid bound while StatNC remains correct.

last, but not least, a validation of the framework in a practical setting like the ones mentioned in the Introduction (MPLS domain, WSN) should provide new insights.

## REFERENCES

[1] D. Anick, D. Mitra, and M. M. Sondhi. Stochastic theory of a data-handling system with multiple sources. *The Bell System Technical Journal*, 61(8):1871–1894, 1982.

[2] D. Awduche, J. Malcom, J. Agogbua, M. O'Dell, and J. McManus. Requirements for traffic engineering over mpls. RFC 2702, September 1999.

[3] G. Bouabene, C. Jelger, C. Tschudin, S. Schmid, A. Keller, and M. May. The autonomic network architecture (ana). *IEEE Journal on Selected Areas in Communications*, 28(1):4–14, 2010.

[4] A. Burchard, J. Liebeherr, and S.D. Patek. A min-plus calculus for end-to-end statistical service guarantees. *IEEE Transactions on Information Theory*, 52(9):4105–4114, 2006.

[5] C.-S. Chang. Stability, queue length and delay of deterministic and stochastic queueing networks. *IEEE Transactions on Automatic Control*, 39(5):913–931, May 1994.

[6] C.-S. Chang. *Performance Guarantees in Communication Networks*. Telecommunication Networks and Computer Systems. Springer-Verlag, 2000.

[7] F. Ciucu, A. Burchard, and J. Liebeherr. A network service curve approach for the stochastic analysis of networks. In *Proc. ACM SIGMETRICS*, pages 279–290, June 2005.

[8] F. Ciucu and J. B. Schmitt. Perspectives on network calculus: no free lunch, but still good value. *Proc. of ACM SIGCOMM*, August 2012.

[9] C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934.

[10] R. L. Cruz. A calculus for network delay, Part I: Network elements in isolation. *IEEE Transactions on Information Theory*, 37(1):114–131, January 1991.

[11] R. L. Cruz. A calculus for network delay, Part II: Network analysis. *IEEE Transactions on Information Theory*, 37(1):132–141, January 1991.

[12] M. Fidler. An end-to-end probabilistic network calculus with moment generating functions. In *Proc. of IEEE IWQoS*, pages 261–270, June 2006.
[13] Y. Gao and Y. Jiang. Analysis on the capacity of a cognitive radio network under delay constraints. *IEICE Transactions on Communications*, E95-B(4):1180 – 1189, April 2012.
[14] R. Gibbens and F. Kelly. Measurement-based connection admission control. In *Proc. 15th International Teletraffic Congress*, pages 781–790, 1997.
[15] M. Grossglauser and D. N.C. Tse. A framework for robust measurement-based admission control. *IEEE/ACM Transactions on Networking*, 7:293–309, June 1999.
[16] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.
[17] S. Jamin, P.B. Danzig, S. J. Shenker, and L. Zhang. A measurement-based admission control algorithm for integrated services packet networks. *IEEE/ACM Transactions on Networking*, 5:56–70, February 1997.
[18] Y. Jiang and Y. Liu. *Stochastic Network Calculus*. Springer, 2008.
[19] J. Kurose. On computing per-session performance bounds in high-speed multi-hop computer networks. In *Proc. of ACM SIGMETRICS*, pages 128–139, 1992.
[20] J. Liebeherr, A. Burchard, and F. Ciucu. Delay bounds in communication networks with heavy-tailed and self-similar traffic. *IEEE Transactions on Information Theory*, 58:1010–1024, 2012.
[21] R. Lübben, M. Fidler, and J. Liebeherr. A foundation for stochastic bandwidth estimation of networks with random service. In *Proc. of IEEE INFOCOM*, pages 1817–1825, 2011.
[22] P. Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *Ann. Probab.*, 18(3):1269–1283, 1990.
[23] J. Qiu and E. W. Knightly. Measurement-based admission control with aggregate traffic envelopes. *IEEE/ACM Transactions on Networking*, 9(2):199–210, April 2001.
[24] A. Rizk and M. Fidler. Non-asymptotic end-to-end performance bounds for networks with long range dependent fbm cross traffic. *Computer Networks*, 56(1):127–141, January 2012.
[25] K. Wang, F. Ciucu, C. Lin, and S. Low. A stochastic power network calculus for integrating renewable energy sources into the power grid. *IEEE Journal on Selected Areas in Communications: Smart Grid Communications Series*, 30:1037–1048, 2012.
[26] O. Yaron and M. Sidi. Performance and stability of communication networks via robust exponential bounds. *IEEE/ACM Transactions on Networking*, 1(3):372–385, June 1993.

## Appendix I

Assume a Markov modulated arrival $A$ with finite signal space $S$. I.e. the distribution of the increments of $A$ depend on the current state of an underlying Markov chain $Y$. The states of this Markov chain are described by the set $S$ and transition matrix $T = [t_{rs}]$, such that $t_{rs} > 0$ for all $(r, s) \in S^2$.[3] Denote the state of the Markov chain at time $n \in \mathbb{N}$ by $y_n$, further denote by $E$ a diagonal matrix with entries $E_s := E_{ss} := \mathbb{E}(e^{\theta a_k} | y_k = s)$ for all $s \in S$. It holds the following:

**Lemma.** *$A$ is $(\sigma(\theta), \rho(\theta))$-bounded with:*

$$\sigma(\theta) = {}^1\!/\!_\theta \log \left( \max_{s \in S} E_s \cdot \frac{\max_{s \in S} x_s}{\min_{s \in S} x_s} \rho(E \cdot T)^{-1} \right)$$

$$\rho(\theta) = {}^1\!/\!_\theta \log \left( \rho(E \cdot T) \right)$$

*where $\rho : Mat(S) \to \mathbb{R}$ is the spectral radius of some matrix over $S$ and the vector $x$ is a positive eigenvector of $ET$.*

*Remark.* The proof given here is very similar to a proof in [6]. The difference is, that we give constructive proof for a $(\sigma(\theta), \rho(\theta))$-bound, while there (Example 7.2.7. together with Lemma 7.2.6.) only the existence of such a bound is proven.

---

[3] We will use the symbol $S$ for both, the set of states, as well as their cardinality. This causes in fact no trouble and is quite intuitive.

*Proof.* Fix $\theta > 0$. We start with a backward-like equation, for every $r \in S$ holds:

$$\begin{aligned}\mathbb{E}(e^{\theta A(n)} | \, y_1 = r) =: E_r(n) &= \sum_{s\in S} \mathbb{E}(e^{\theta A(n)} | \, y_1 = r, \, y_2 = s)\mathbb{P}(y_2 = r | \, y_1 = s) \\ &= \sum_{s\in S} \mathbb{E}(e^{\theta a_1} | \, y_1 = r, \, y_2 = s)\mathbb{E}(e^{\theta A(n)-a_1} | \, y_1 = r, \, y_2 = s)t_{rs} \\ &= E_r \cdot \sum_{s\in S} \mathbb{E}(e^{\theta A(n-1)} | \, y_1 = s)t_{rs} \\ &= E_r \cdot \sum_{s\in S} E_s(n-1)t_{rs}\end{aligned}$$

If we denote by $E(n)$ the vector with entries $[E_r(n)]$, we can write the above short by:

$$E(n) = ET \cdot E(n-1)$$

Using this recursion we get:

$$E(n) = (ET)^{n-1} E \cdot \mathbf{1},$$

where $\mathbf{1}$ is the unit column vector on $S$. Assume now the beginning state of the chain is not given, but follows an (arbitrary) distribution $\pi = (\pi_1, \ldots, \pi_S)$, then an application of the law of total probability yields:

$$\begin{aligned}\mathbb{E}_\pi(e^{\theta A(n)}) &= \sum_{s\in S} \mathbb{P}(y_1 = s)E_s(n) \\ &= \sum_{s\in S} \pi_s E_s(n) \\ &= \sum_{s\in S} \pi_s \left((ET)^{n-1} E \cdot \mathbf{1}\right)_s \\ &= \pi \cdot (ET)^{n-1} E \cdot \mathbf{1}\end{aligned}$$

Next we want to bound the entries in the matrix $(ET)^{n-1}$. The following corollary (8.1.33. in [16]), achieves that: □

**Corollary.** *Let $A = [a_{rs}] \in Mat(S)$ be a nonnegative matrix. Write $A^m = [a_{rs}^{(m)}]$. If $A$ has a positive eigenvector $x = [x_s]$ then for all $m \in \mathbb{N}$ and all $r \in S$ holds:*

$$\sum_{s\in S} a_{rs}^{(m)} \leq \frac{\max_{t\in S} x_t}{\min_{t\in S} x_t} \cdot \rho(A)^m$$

*Proof. (continuation of proof)* To apply this corollary we need a positive eigenvector $x$. Since $T$ is positive and $E$ has positive entries on the diagonal, we know that also the matrix $ET$ is positive (i.e. every entry is larger 0) and hence also $(ET)^n$for every $n \in \mathbb{N}$. This allows us to apply the Perron-Frobenius theorem, which guarantees an eigenvector with positive entries. Denote this eigenvector by $x = [x_s] \in \mathbb{R}^S$.

Eventually we have for every starting distribution $\pi$ and all $n \in \mathbb{N}$:

$$\begin{aligned} E_\pi(n) &= \sum_{r\in S}\pi_r\sum_{s\in S}(ET)^{n-1}_{rs}(E\cdot\mathbf{1})_s \\ &\le \sum_{r\in S}\pi_r\left(\max_{t\in S}E_t\right)\sum_{s\in S}(ET)^{n-1}_{rs} \\ &\le \left(\max_{t\in S}E_t\right)\sum_{r\in S}\pi_r\cdot\frac{\max_{t\in S}x_t}{\min_{t\in S}x_t}\rho(ET)^{n-1} \\ &= \left(\max_{t\in S}E_t\right)\cdot\frac{\max_{t\in S}x_t}{\min_{t\in S}x_t}\rho(ET)^{n-1} \end{aligned}$$

Note that the given bound does not depend on the initial distribution $\pi$. This allows us to finish the proof by using:

$$\mathbb{E}(e^{\theta A(m,n)}) = \mathbb{E}(e^{\theta(A(n)-A(m)}|\,Y_m) = E_{Y_m}(n-m)$$

□

## Appendix II

The expectation of an exponentially distributed random variable with parameter $\lambda$ and truncated by $M$ is given by $1/\lambda(1-e^{-\lambda M})$. Further the expectation of a truncated Pareto distribution with parameters $x_{min}$ and $s \neq 1$, is given by

$$x_{min} + \frac{x_{min}}{1-s}\cdot\left(\left(\frac{x_{min}}{M}\right)^{s-1}-1\right)$$

and for $s = 1$ by:

$$x_{min} + x_{min}\log\left(\frac{M}{x_{min}}\right)$$

The fitting $\lambda$ is found by solving

$$\frac{1}{\lambda}\left(1-e^{-\lambda M}\right) \stackrel{!}{=} x_{min} + \frac{x_{min}}{1-s}\cdot\left(\left(\frac{x_{min}}{M}\right)^{s-1}-1\right)$$

or

$$\frac{1}{\lambda}\left(1-e^{-\lambda M}\right) \stackrel{!}{=} x_{min} + x_{min}\log\left(\frac{M}{x_{min}}\right)$$

numerically.

## Appendix III

We give in this appendix some insights about the fact, that $\alpha$ can be choosen rather freely, without worsening the bounds much or stated differently: optimization over $\alpha$ does not improve the bounds by any significant value. We will show this for the non-parametric estimator of Subsection 4.2. First we rewrite the bound of 5 as a function of $\alpha$, remember we search for the smallest $b$ such that $\mathbb{P}(q(n) > b) \le \varepsilon$ holds. Insterting the statistical bound and resolving for $b$ we get:

$$b(\alpha) = -\frac{1}{\theta}\log\left(\varepsilon(1-e^{-\theta c}\bar{A}) - \alpha(1-e^{-\theta c}\bar{A}) - q(\alpha)(e^{\theta M}-1)e^{-\theta c}\right)$$

with $q(\alpha) := (\varepsilon-\alpha)\sqrt{\frac{-\log(\alpha/2)}{|n_0|}}$. We investigate now the function $q$. It is easy to see, that its range lies in $(0,\infty)$ with diverging to infitiy while $\alpha \to 0$ and zero as

| $\alpha$ | $10^{-10}$ | $10^{-100}$ | $10^{-1000}$ | $10^{-10000}$ |
|---|---|---|---|---|
| $q(\alpha)$ | $\approx 1.5 \cdot 10^{-5}$ | $\approx 4 \cdot 10^{-5}$ | $\approx 1.5 \cdot 10^{-4}$ | $\approx 4 \cdot 10^{-4}$ |

TABLE 1. Values for $q(\alpha)$ with $\varepsilon = 10^{-4}$ and $|n_0|$.

limit for $\alpha \to \varepsilon$. Due to the logarithm under the squareroot we have for $q(\alpha)$ the property that it remains relatively unchanged for most choices of $\alpha$. The following table (with an $\varepsilon = 10^{-4}$) makes this clear. Please note that values of $\alpha$ have been choosen, which lie far beyond the numerical precision of usual computers:

Esssentially this allows us to bound the value of $q(\alpha) \leq Q$ from above (in our example $Q = 4 \cdot 10^{-4}$ could be a bound) and reduce the contribution of the summand $\alpha(1 - e^{-\theta c}\bar{A})$ at the same time to a diminishing value (by choosing $\alpha$very small). This leaves us with a bound, which is determined effectively by:

$$b_Q \approx -\frac{1}{\theta} \log \left( \varepsilon(1 - e^{-\theta c}\bar{A}) - Q(e^{\theta M} - 1)e^{-\theta c} \right)$$

Although this is a special case, we have encountered this "independence" from $\alpha$ through all our experiments.