

Window Flow Control in Stochastic Network Calculus

Michael A. Beck and Jens Schmitt

Technical Report (391/15)

University of Kaiserslautern, Germany.

Distributed Computer Systems Lab (DISCO)

{beck}{jschmitt}@cs.uni-kl.de

ABSTRACT. Feedback is omnipresent in communication networks. One prominent example is window flow control (WFC) as, e.g., found in many transport protocols, for instance TCP. In deterministic network calculus elegant closed-form solutions have been derived to provide performance bounds for WFC systems. However, a treatment of WFC in stochastic network calculus (SNC) has so far been elusive. In this work, we present the first WFC analysis in SNC for subadditive and general service in the feedback loop. The subadditive case turns out as an application of existing results, switching to continuous time requires more effort. We further discuss how the condition of subadditivity is preserved under concatenation of servers and demultiplexing of flows. The key idea for the general case is to keep track of how much the service deviates from being subadditive. Both methods are illustrated in numerical examples and their properties are discussed.

CHAPTER 1

Introduction

Stochastic Network Calculus (SNC) has matured in recent years to provide an alternative method for performance analysis of stochastic queueing systems (see e.g. [14, 12, 8]). Many results from the deterministic network calculus (DNC) have been transferred into the stochastic domain, some have been rather immediate some have required considerable effort (e.g., deriving the end-to-end service [6]). One major remaining open issue is the stochastic analysis of feedback-based systems, such as Window Flow Controlled (WFC) transport protocols, e.g. TCP. While there are very elegant solutions for WFC in the deterministic setting [1, 4, 15], WFC in SNC has been identified for some time already as a very challenging open research question [14, 11, 12, 7]. Moreover, being able to analyse WFC systems in SNC would be very relevant to open up new application areas for SNC such as modelling smart grid systems [13], for instance.

In this work, we present an approach to analyse WFC in SNC in two ways. The first works under a subadditivity assumption for the involved service processes, while the second works without such an assumption.

First we tackle the WFC in SNC by restricting to a subadditive service provided by the network which is to be flow controlled. This restriction is motivated by the crucial role the so-called subadditive closure plays in the WFC analysis. Assuming the network service to be subadditive eases the computation of the subadditive closure tremendously. While the subadditivity assumption is clearly restrictive, we argue it is a reasonable first step and, as we discuss in Section 5.3, it can be met in a relevant class of applications.

Under the subadditive service assumption, we analyse WFC in SNC for both, discrete and continuous time models. While in discrete time the analysis runs smoothly and, in fact, requires mainly an application of existing results, the continuous time model takes more effort and care, mainly due to subtleties introduced by instantaneous bursts from cross-traffic arrivals. Nevertheless, we present closed-form solutions in both cases for WFC within the so-called MGF calculus [5, 10], a sub-branch of the SNC.

The key idea for the general case is to stochastically control how far the service deviates from being subadditive and cast this into the setting of MGF-calculus, as further contribution to the violation probabilities of the performance bounds. We demonstrate the method by providing some numerical results in case of a server in the feedback loop that is not subadditive.

CHAPTER 2

Notations and Basic Results

We give here the needed notations and basic results of network calculus; for further details see [5, 3]. We start by defining flows which enter and depart service elements.

DEFINITION 1. We denote a *flow* by its cumulatives A , i.e., $A(t)$ counts data arrivals in the interval $[0, t]$. The bivariate extension of A is defined by $A(s, t) := A(t) - A(s)$. In case of discrete time we also consider the *increments* $a(t) = A(t) - A(t - 1)$. In case of continuous time we assume A to be cadlag (right continuous, left limits).

We introduce two service descriptions here, one for the univariate calculus (usually used in DNC and the tail-bound-branch of SNC) and the bivariate calculus (used in the MGF-branch of SNC):

DEFINITION 2. We say a service element offers a *service curve* U , if for any input-output pair A, B and time t

$$B(t) \geq A \otimes U(t) := \min_{0 \leq s \leq t} \{A(s) + U(t - s)\}$$

(in continuous time the min is to be replaced by inf).

Let U be a bivariate function with $U(s, t) \leq U(s, t')$ for all $t \leq t'$ and time be discrete. A service element is a *dynamic U -server*, if for any input-output pair A, B and time t :

$$B(t) \geq A \otimes U(0, t) := \min_{0 \leq s \leq t} \{A(0, s) + U(s, t)\}.$$

The operator \otimes is called min-plus convolution. Note that the bivariate U is in general not a flow, i.e., there may exist r, s, t with $U(s, t) \neq U(s, r) + U(r, t)$.

The following two results enable the analysis of feedforward networks (under arbitrary multiplexing).

THEOREM 3. *Consider two service elements, such that the output of the first service element is the input to the second. If both service elements have a service curve U_i (are dynamic U_i -server, $i = 1, 2$), the system offers a service curve $U_1 \otimes U_2$ (is a dynamic $(U_1 \otimes U_2)$ -server).*

THEOREM 4. *Consider a service element serving two flows A_1, A_2 . If the server offers a service curve U_o and the flow A_1 is bounded by $A_1(s, t) \leq \alpha(t - s)$, then the system offers a service curve $U(t) = U_o(t) - \alpha(t)$ to flow A_2 .*

For discrete time, if the element is a dynamic U_o -server, the system acts as a dynamic U -server with $U(s, t) = U_o(s, t) - A_1(s, t)$ for flow A_2 .

The third network-operation we present here is central to this work. It describes, how a feedback system as presented in Figure 2.0.1 is handled. In this

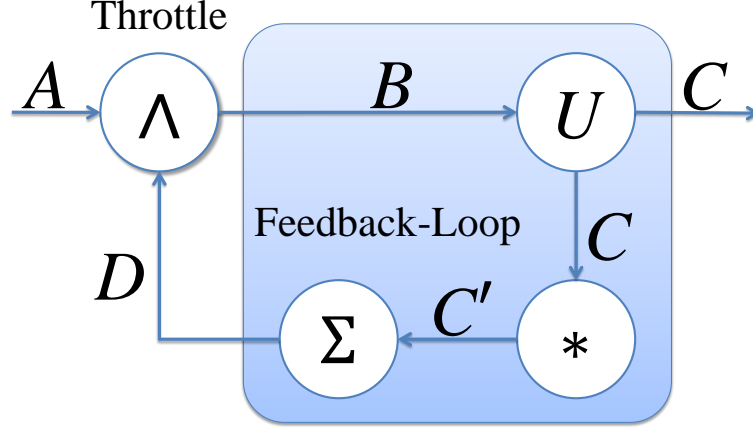


FIGURE 2.0.1. A window flow controller: the input A is throttled at the \wedge -element. The departures of the system are C and the feedback-loop consists of the dynamic U -server, an unspecified service element and a window-element. The $*$ is a placeholder for zero, one, or several elements, like delay-elements, scalars or dynamic servers.

feedback system, the original input A is fed to a throttle-element first, which governs, how much data is admitted to the system (the feedback-loop). This is realized by taking the minimum of the input A and the output of the service element at any time, such that:

$$B(t) = A(t) \wedge D(t).$$

Such systems are studied for example in [1]. We formulate from there the following theorem:

THEOREM 5. *Let time be discrete. Assume the whole feed-back-loop in Figure 2.0.1 is described by a service curve U_{fb} (is a dynamic U_{fb} -server). The throttle element \wedge has a service curve U_{\wedge} (is a dynamic U_{\wedge} -server), with:*

$$U_{\wedge}(t) = \bigwedge_{k=0}^{\infty} U_{fb}^{(k)}(t) \quad \left(U_{\wedge}(s, t) = \bigwedge_{k=0}^{\infty} U_{fb}^{(k)}(s, t) \right).$$

Here, the notation $U_{fb}^{(k)}$ stands for the k -fold self-convolution of U_{fb} . Further, for any U we define $U^{(0)}$ to be the neutral element of the convolution: $U^{(0)}(t) = \mathbf{1}(t) = \infty$ for all $t > 0$ and $\mathbf{1}(0) = 0$ (for the bivariate case we have $\mathbf{1}(s, t) = \infty$ for all $s < t$ and $\mathbf{1}(t, t) = 0$ for all t , respectively). The expression $\bigwedge_{k=0}^{\infty} U^{(k)} =: \bar{U}$ is known as the *subadditive closure* of service U . Subadditivity in the context of

univariate and bivariate functions means for all times r, s, t :

$$\begin{aligned} U(t) &\leq U(s) + U(t-s) \\ U(s, t) &\leq U(s, r) + U(r, t) \end{aligned}$$

As the name suggests the subadditive closure of a service description is subadditive. Further, for any subadditive U we have $\bigwedge_{k=0}^{\infty} U^{(k)} = \mathbf{1} \wedge U$.

To complete the analysis of the feedback system in Figure 2.0.1 we combine Theorem 5 and Theorem 3 for a service description of the whole system:

$$U_{sys} = U_{\wedge} \otimes U.$$

This service description can then be used to derive probabilistic performance bounds, e.g., on end-to-end-delay, as detailed below.

In network calculus, one assumes systems to be empty at time zero, i.e., $A(0) = 0$. To avoid trivial cases in the feedback system one needs a window element in the feedback-loop. It serves two purposes: First it kickstarts our system by initially admitting a certain amount of data to U , and, second, it will control how much data is inside the feedback-loop, thus guaranteeing a maximal backlog on any element inside of it. We define such a window element by:

DEFINITION 6. A *window element* is described by a cadlag function from $\Sigma : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ (or in discrete time just by a sequence $\Sigma : \mathbb{N}_0 \rightarrow \mathbb{R}_0^+$), such that for any input-flow A it produces an output B by:

$$(2.0.1) \quad \begin{aligned} B(t) &= A(t) + \Sigma(t) \\ \Sigma(s, t) &\geq -A(s, t) \quad \forall s \leq t. \end{aligned}$$

We further define the bivariates $\Sigma(s, t)$ as for flows.

The constraint (2.0.1) enforces that the output $B(s, t)$ of a window element cannot be negative.

Using the above calculus one can derive probabilistic performance bounds. For this we define two performance measures:

DEFINITION 7. The *backlog* q of the system at time t is defined by

$$q(t) := A(t) - B(t).$$

The *virtual delay* d at time t is defined by

$$d(t) := \inf\{T : A(t) \leq B(t+T)\}.$$

To achieve actual bounds on q and d we need to add a probabilistic component. There are different methods [12] to describe the stochastic nature of flows and service descriptions, one of them is [5, 10]:

THEOREM 8. Let time be discrete and A and the service process U be stochastically independent. Let $\theta > 0$ and:

$$(2.0.2) \quad \mathbb{E}(e^{\theta A(s,t)}) \leq e^{\theta \rho_A(\theta)(t-s) + \theta \sigma_A(\theta)}$$

$$(2.0.3) \quad \mathbb{E}(e^{-\theta U(s,t)}) \leq e^{\theta \rho_U(\theta)(t-s) + \theta \sigma_U(\theta)}$$

for all $s \leq t$ and $\rho_A(\theta) + \rho_U(\theta) := \rho(\theta) < 0$. Then:

$$\begin{aligned} \mathbb{P}(q(t) > x) &\leq e^{\theta(\sigma_A(\theta) + \sigma_U(\theta) - x)} (1 - e^{\theta \rho(\theta)})^{-1} \quad \forall t \geq 0, \\ \mathbb{P}(d(t) > T) &\leq e^{\theta(\sigma_A(\theta) + \sigma_U(\theta) + \rho_U(\theta)T)} (1 - e^{\theta \rho(\theta)})^{-1} \quad \forall t \geq 0 \end{aligned}$$

The condition $\rho(\theta) < 0$ is a stability condition for the system. There are versions of the above theorem for continuous time, as well as for the case in which A and U are not stochastically independent [2]. The above bounds use Moment-Generating Functions (MGFs) to describe the behaviour of A and U , which is why we just call it MGF-calculus for short.

We point out here that none of the state-of-the-art methods is able to describe the service process resulting from a feedback-system (Theorem 5). Thus, the solution of WFC systems eluded stochastic network calculus, so far.

CHAPTER 3

Problem Exposition

Subadditivity plays an important role when analysing WFC systems, as the throttle's service is just the subadditive closure of the elements inside the feedback loop. Hence, feedback-loops containing subadditive elements are much easier to analyse. We present this situation now in detail.

We look at a feedback loop which contains a subadditive service element U (lying on the path from the sender to the receiver) and a window-element (lying on the return path). Such a system can be easily analysed, by applying the most general result concerning WFC systems as found in the textbook of Chang [5]. To that end, denote the space of bivariate functions which are monotonically increasing in their second variable as $\tilde{\mathcal{F}}$. We see immediately that all flows and all dynamic-server descriptions lie in $\tilde{\mathcal{F}}$, as they fulfill $F(s, t) \leq F(s, t')$ for all $t \leq t'$.

DEFINITION 9. An operator $\pi : \tilde{\mathcal{F}} \rightarrow \tilde{\mathcal{F}}$ is called σ -additive if:

$$(3.0.4) \quad \pi \left(\bigwedge_{n=1}^{\infty} F_n \right) = \bigwedge_{n=1}^{\infty} \pi(F_n),$$

where F_n is any sequence in $\tilde{\mathcal{F}}$ and the infima are understood pointwisely.

One can easily verify that the space of σ -additive operators is closed under taking countable minima and compositions. Further they distribute over countable minima:

$$\left(\bigwedge_{n=1}^{\infty} \pi_n \right) \circ \pi = \bigwedge_{n=1}^{\infty} (\pi_n \circ \pi), \quad \pi \circ \left(\bigwedge_{n=1}^{\infty} \pi_n \right) = \bigwedge_{n=1}^{\infty} (\pi \circ \pi_n)$$

An example of a σ -additive operator is the convolution with some bivariate $U \in \tilde{\mathcal{F}}$, defined by $\pi_U(A) := A \otimes U$. Other examples needed for our specific feedback system are

$$\begin{aligned} \pi_{\tilde{e}}(A) &:= A \\ \pi_{+w}(A) &:= \begin{cases} A(s, t) & \text{if } s \neq 0 \\ A(0, t) + w & \text{if } s = 0 \end{cases} \end{aligned}$$

where $w \in \mathbb{R}$. The above operators represent the identity operator and a window-element, respectively; their σ -additivity as well as the following properties are easy to verify:

- π_{+w} commutes with π_U , i.e. $\pi_U \circ \pi_{+w} = \pi_{+w} \circ \pi_U$.
- $\pi_{+w}^2 = \pi_{+2w}$, and U is subadditive iff π_U is idempotent, i.e. $\pi_U^2 = \pi_U$.

The whole feedback loop can be expressed by successively applying σ -additive operators and it is a σ -additive operator itself:

$$\pi_{fb} := \pi_{+w} \circ \pi_U,$$

i.e. $D = \pi_{fb}(B)$. The relations between the flows in Figure 2.0.1 are hence given by:

$$(3.0.5) \quad B = A \wedge \pi_{fb}(B).$$

Lemma 5.7.2. in [5] can be applied resulting in $B \geq \pi^*(A)$ for any B fulfilling $B \geq A \wedge \pi_{fb}(B)$, where $\pi^* = \pi_{\bar{e}} \wedge \bigwedge_{n=1}^{\infty} \pi^n$ is the closure of π .

We show now that π^* is tractable when considering a subadditive U together with a window-element.

EXAMPLE 10. Applying the above properties to π_{fb} , we observe that

$$\pi_{fb}^n \geq \pi_{+nw} \circ \pi_U.$$

And for the closure of π_{fb}

$$\pi_{fb}^* \geq \pi_{\bar{e}} \wedge \bigwedge_{n=1}^{\infty} \pi_{+nw} \circ \pi_U = \pi_{\bar{e}} \wedge \pi_U.$$

Applying the result of Chang we have for the departures C of the feedback system:

$$\begin{aligned} C(t) &\geq \pi_U(B)(0, t) \geq \pi_U \circ \pi_{fb}^*(A)(0, t) \\ &= \bigwedge_{n=0}^{\infty} \pi_U \circ \pi_{fb}^n(A)(0, t) \\ &\geq \bigwedge_{n=0}^{\infty} \pi_{+nw} \circ \pi_U(A)(0, t) = \pi_U(A)(0, t) \end{aligned}$$

We use the distributivity of σ -additive operators in the second line. We see that the whole system behaves, just like the unthrottled one. With given MGF-bounds on A and U we can use this bound on C to produce stochastic performance bounds, as in Theorem 8.

The above example makes the role of subadditivity clear: without U being subadditive, we would not have $\pi_U^2 = \pi_U$ and the description of π_{fb}^* would include π_U^n . The repeated application of π_U prevents deriving stochastic performance bounds as above. Note further, that the window-element causes only minor difficulties, as its operator commutes with π_U .

The most important example for a non-subadditive U is a service, which resulted by applying Theorem 3. How one can preserve subadditivity instead for a concatenation of service elements is discussed in Section 5.3. However, preserving subadditivity in that way comes with the cost of a decreased service.

We now leave the notations of σ -additive operators, and take a step back to have a closer look at subadditivity in the most simple scenario.

There are some crucial differences for a feedback system, when switching from univariate to bivariate descriptions. Most of them stem from the different service characterization. For this work, the most important difference is found in Theorem 4. Note that the leftover service descriptions differ by using the function α :

$$(3.0.6) \quad U(t) = U_o(t) - \alpha(t)$$

$$(3.0.7) \quad U(s, t) = U_o(s, t) - A_1(s, t).$$

Calculating the self-convolution of both service descriptions reveals the importance of that difference. If we assume for a while U_o to be subadditive, we obtain:

$$\begin{aligned} U \otimes U(t) &= \inf_{0 \leq s \leq t} \{U_o(s) - \alpha(s) + U_o(t-s) - \alpha(t-s)\} \\ &\geq U_o(t) - \sup_{0 \leq s \leq t} \{\alpha(s) + \alpha(t-s)\}. \end{aligned}$$

And, as the last supremum is in general larger than $\alpha(t)$, we *cannot* achieve subadditivity of U . Consequently, the subadditive closure \bar{U} becomes non-trivial.

In contrast to that, (3.0.7) yields

$$\begin{aligned} &U \otimes U(s, t) \\ &= \inf_{s \leq r \leq t} \{U_o(r-s) - A_1(s, r) + U_o(t-r) - A_1(r, t)\} \\ (3.0.8) \quad &\geq U_o(t-s) - A_1(s, t) = U(s, t) \end{aligned}$$

and hence U is subadditive, resulting in $\bar{U} = \mathbf{1} \wedge U$. This gives us an important case at hand which is simple to analyse. The condition of U_o being subadditive is restrictive, though, and not always given. The most prominent example of a service which is not subadditive is the result of applying Theorem 3. We explore the space of subadditive service in Section 5.3 in detail.

As the subadditive closure of the feedback-loop appears in the solution of Theorem 5, subadditivity of the service curve U and how much is “missing” to it are crucial questions in WFC systems. Assume we can give the following bound, describing by what quantity a service curve U fails subadditivity:

$$(3.0.9) \quad U(t) - U \otimes U(t) \leq b \quad \forall t \in \mathbb{R}_0^+.$$

Or, written differently:

$$U^{(2)}(t) \geq U(t) - b \quad \forall t \in \mathbb{R}_0^+.$$

If we consider some U and b to fulfill Equation (3.0.9), we achieve for $\Sigma \geq b$ the subadditivity of $U + \Sigma = U_{fb}$, since

$$\begin{aligned} (U + \Sigma) \otimes (U + \Sigma)(t) &= U^{(2)}(t) + 2\Sigma \\ &\geq U(t) - b + 2\Sigma \\ &\geq U(t) + \Sigma = U_{fb}(t). \end{aligned}$$

A subadditive U_{fb} , however means that¹:

$$U_{sys}(t) = \bar{U}_{fb} \otimes U(t) = (\mathbf{1} \wedge U + \Sigma) \otimes U(t) = U(t).$$

This effect, that a system with a large enough window can be analysed as unthrottled, is well known and can for example be found in [1]. One can think of the window-element to compensate for how the service deviates from subadditivity. An approach following this idea is presented in Chapter 6.

¹As they are σ -additive operators the convolution with a function is distributive over minima, i.e.: $(U_1 \wedge U_2) \otimes V = U_1 \otimes V \wedge U_2 \otimes V$

CHAPTER 4

Bivariate Continuous Time Equations

Theorem 5 only applies for discrete time in the bivariate case. In *continuous* time, however, literature provides no bivariate service description for the throttle. Closing that gap proves to be a rather technical problem, including many continuity arguments. For ease of notation, we leave the placeholder in Figure 2.0.1 empty, throughout this chapter.

To deal with continuous time we need to consider left-limits of flows:

For a flow A we define the left-limits $A^\circ(t) := \lim_{r \nearrow t} A(r)$ and the bivariate notation:

$$\tilde{A}(s, t) := A(t) - \lim_{r \nearrow s} A(r).$$

While not being a flow, we use the notation $\tilde{\Sigma}$ and Σ° for window elements correspondingly.

Due to allowing instantaneous bursts in arrivals A we need to redefine a dynamic U -server.

DEFINITION 11. In continuous time a *dynamic U -server* satisfies for any input-output pair A and B :

$$B(t) \geq A^\circ \otimes U(0, t) = \inf_{0 \leq s \leq t} \{A^\circ(s) + U(s, t)\}.$$

Further we call a dynamic U -server *proper*, if there exists some Lipschitz-constant $r_{max} \in \mathbb{R}_0^+$, such that the output is Lipschitz-continuous, i.e. $B(s, t) \leq r_{max}(t - s)$ for all $s \leq t$.

The difference between Definition 2 and the above is to replace A by its left-limits. We now give some intuition for this: let time s be the beginning of a backlogged period with A having a burst of size b , i.e., $\tilde{A}(s, s) = b$. Due to the right-continuity of A we only have: $B^\circ(s) = A^\circ(s)$ and $B(s) < A(s)$. This circumstance prevents us from formulating a relationship between B and A like $B(t) = A(s) + U(s, t)$, and hence achieving $B(t) \geq A \otimes U(0, t)$ is out of reach.

Fortunately, this slight difference in the service description extends naturally to the usual results of network calculus. The next lemma illustrates the influence of the modified service description.

LEMMA 12. *Assume the situation as in Theorem 4, but for continuous time. The system is a dynamic U -server for A_2 with*

$$U = U_o - \tilde{A}_1.$$

Further, if $U_o(\cdot, \cdot)$ is proper, then $U(\cdot, \cdot)$ is also proper.

With the definitions of flows and service elements in place we can revisit the feedback system.

For the departures B of the throttle and all $s \leq t$ we have:

$$\begin{aligned} B(s, t) &= A(s, t) + q_\wedge(s) \wedge C(s, t) + \lim_{r \nearrow s} \Sigma(r) - q_U(s) + \tilde{\Sigma}(s, t) \\ &= A(t) - B(s) \wedge C(s, t) + \Sigma(t) - q_U(s) \\ &= A(t) - B(s) \wedge C(t) - B(s) + \Sigma(t), \end{aligned}$$

with $q_\wedge(s)$, $q_U(s)$ denoting the backlog at the throttle and the dynamic U -server at time s , respectively. The second line reads as follows: the newly admitted flow in the interval $(s, t]$ is the minimum of two expressions: 1) the amount of data arriving by $A(s, t)$, plus the already accumulated data at the throttle at time s ; 2) the amount of data being delivered $C(s, t)$, plus the difference between allowed data in the system and already queued data at U , plus any change in the window $\tilde{\Sigma}(s, t)$.

We use the service description of U to replace $C(t)$ and get:

$$B(s, t) \geq A(t) - B(s) \wedge B^\circ \otimes U(0, t) - B(s) + \Sigma(t).$$

Adding $B(s)$ gives for all $t \geq 0$:

$$(4.0.10) \quad B(t) \geq A(t) \wedge B^\circ \otimes U(0, t) + \Sigma(t).$$

Now we ask which flows B fulfill (4.0.10). For this we need the window element to be Lipschitz-continuous with Lipschitz-constant σ :

$$(4.0.11) \quad \Sigma(s, t) \leq \sigma(t - s) \quad \forall s \leq t$$

THEOREM 13. *Let U be proper and non-negative, Σ fulfill (4.0.11) and $\Sigma(t) \geq \Sigma_{min} > 0$ for all $t \in \mathbb{R}_0^+$. Then for all arrivals A exists a unique B' , fulfilling (4.0.10) with equality and if some B fulfills (4.0.10), also $B \geq B'$ holds. Further B' is pointwisely given by:*

$$\begin{aligned} B'(t) &= \lim_{i \rightarrow \infty} B^{(i)}(t) \quad \forall t \geq 0 \quad B^{(0)}(t) = A(t) \\ B^{(i)}(t) &= A(t) \wedge (B^{(i-1)})^\circ \otimes U(0, t) + \Sigma(t) \end{aligned}$$

PROOF. This proof is very similar to the one in [1]. However, there are a few obstacles to circumvent, hence we give the complete proof again.

Denote by \mathcal{B} the set of flows satisfying (4.0.10). Since $A \in \mathcal{B}$, we know $\mathcal{B} \neq \emptyset$. We can henceforth define pointwisely:

$$B'(t) = \inf_{B \in \mathcal{B}} B(t)$$

We claim that B' is a flow and again fulfills (4.0.10). We start with showing B' is a flow, in expression it is non-decreasing and right-continuous. For the first let $s < t$ be arbitrary and let $B_k \in \mathcal{B}$ be a sequence of flows such that $\lim_{k \rightarrow \infty} B_k(t) = B'(t)$, then

$$B'(t) - B'(s) = \lim_{k \rightarrow \infty} B_k(t) - B'(s) \geq \lim_{k \rightarrow \infty} B_k(t) - \liminf_{k \rightarrow \infty} B_k(s) \geq 0,$$

since all B_k are flows and B' is non-decreasing. Now let $t_0 > 0$ be arbitrary. We can find a $B \in \mathcal{B}$ with

$$B(t_0) - B'(t_0) < \delta/2$$

and an $\varepsilon > 0$, such that

$$B(t') - B(t_0) < \delta/2$$

for all $t' \in [t_0, t_0 + \varepsilon)$ (such an ε exists, as B is a flow and hence right-continuous). Adding above two inequalities leads to

$$\delta > B(t') - B'(t_0) \geq B'(t') - B'(t_0)$$

for all $t' \in [t_0, t_0 + \varepsilon)$. Hence B' is right-continuous and henceforth a flow.

Now we show B' fulfills (4.0.10): let t be arbitrary and assume first, that $\inf_{B \in \mathcal{B}} B(t) = \min_{B \in \mathcal{B}} B(t)$ and let $B^* \in \mathcal{B}$, such that $B^*(t) = B'(t)$. Then:

$$B'(t) = B^*(t) \geq A(t) \wedge (B^*)^\circ \otimes U(0, t) + \Sigma(t) \geq A(t) \wedge (B')^\circ \otimes U(0, t) + \Sigma(t)$$

where we have used the monotonicity of \otimes . Now assume the infimum is not adopted by a flow in \mathcal{B} . Denote by B_k a sequence of flows in \mathcal{B} , such that their limit at time t equals $B'(t)$. Note that $\liminf_{k \rightarrow \infty} B_k(s) \geq B'(s)$ for all $s < t$. We have then:

$$\begin{aligned} B'(t) &= \lim_{k \rightarrow \infty} B_k(t) \geq \liminf_{k \rightarrow \infty} \{A(t) \wedge (B_k)^\circ \otimes U(0, t) + \Sigma(t)\} \\ &= A(t) \wedge \liminf_{k \rightarrow \infty} \{(B_k)^\circ \otimes U(0, t)\} + \Sigma(t) \\ &\geq A(t) \wedge \inf_{0 \leq r \leq t} \{(B')^\circ(r) + U(r, t)\} + \Sigma(t) \\ &= A(t) \wedge (B')^\circ \otimes U(0, t) + \Sigma(t) \end{aligned}$$

Hence we have that B' fulfills (4.0.10).

Next we show, that B' fulfills (4.0.10) with equality. Suppose that $B'(t) > A(t) \wedge (B')^\circ \otimes U(0, t) + \Sigma(t)$ holds. Define $B''(t) := A(t) \wedge (B')^\circ \otimes U(0, t) + \Sigma(t)$ and $B''(s) = B'(s)$ for all $s \neq t$. From our construction we immediately have

$$B'(t) > A(t) \wedge B' \otimes U(0, t) + \Sigma(t) = B''(t)$$

and hence:

$$B''(t) = A(t) \wedge B' \otimes U(0, t) + \Sigma(t) \geq A(t) \wedge B'' \otimes U(0, t) + \Sigma(t)$$

so that $B'' \in \mathcal{B}$, which contradicts the minimality of B' . So B' has to fulfill (4.0.10) with equality.

To prove the first sentence of the theorem all left to show is the uniqueness of the solution B' . Let B^* be another flow fulfilling (4.0.10) with equality, define $t_0 = \inf\{t \in \mathbb{R}_0^+ : B'(t) \neq B^*(t)\} < \infty$. By right-continuity of B' and B^* , we can find for each δ an $\varepsilon > 0$ such that $B'(t) < B'(t_0) + \delta$ and $B^*(t) < B^*(t_0) + \delta$ for all $t \in [t_0, t_0 + \varepsilon)$. Choose $\delta = \Sigma_{min}$. We have then:

$$\begin{aligned} B'(t) &= A(t) \wedge (B')^\circ \otimes U(0, t) + \Sigma(t) \\ &= A(t) \wedge \left(\left(\inf_{0 \leq s \leq t_0} \{(B')^\circ(s) + U(s, t)\} \wedge \inf_{t_0 < s \leq t} \{(B')^\circ(s) + U(s, t)\} \right) + \Sigma(t) \right) \end{aligned}$$

Assume the second inf would minimize the expression in the bracket. Since $(B')^\circ(s) \geq B'(t_0)$ and $U(s, t) \geq 0$ for all $s \in (t_0, t]$, this would result in

$$B'(t) \geq A(t) \wedge B'(t_0) + \Sigma(t) \geq A(t) \wedge B'(t_0) + \Sigma_{min} > A(t) \wedge B'(t),$$

which is only possible for $B'(t) > A(t)$, a contradiction. Hence the first infimum must minimize the expression in the bracket. But in this case - applying the same arguments to B^* we get:

$$B'(t) = A(t) \wedge \left(\inf_{0 \leq s \leq t_0} \{(B')^\circ(s) + U(s, t)\} + \Sigma(t) \right) = B^*(t)$$

This however contradicts the construction of t_0 . Hence B' must be unique.

To prove the second part of the theorem we need to ensure first, that the limit of $B^{(i)}(t)$ exists for all $t \in \mathbb{R}_0^+$. First note that $B^{(i)}(t) \geq 0$ holds for all $i, t \in \mathbb{R}_0^+$. Hence, for showing the existence of the limit it is sufficient to prove $B^{(i)}(t) \leq B^{(i-1)}(t)$ for all $t \in \mathbb{R}_0^+$. Assume this holds for some $i \in \mathbb{N}_0$, by the monotonicity of \otimes we have¹:

$$\begin{aligned} B^{(i+1)}(t) &= A(t) \wedge (B^{(i)})^\circ \otimes U(0, t) + \Sigma(t) \\ &\leq A(t) \wedge (B^{(i-1)})^\circ \otimes U(0, t) + \Sigma(t) = B^{(i)}(t) \end{aligned}$$

together with $B^{(1)}(t) \leq A(t)$ we have that all limits $B^{(i)}(t)$ exists. If we can show, that $\lim_{i \rightarrow \infty} B^{(i)}$ fulfills (4.0.10) with equality the result follows from the already proven uniqueness:

$$\begin{aligned} \lim_{i \rightarrow \infty} B^{(i)}(t) &= \lim_{i \rightarrow \infty} B^{(i+1)}(t) \\ &= \lim_{i \rightarrow \infty} \{A(t) \wedge (B^{(i)})^\circ \otimes U(0, t) + \Sigma(t)\} \\ &= A(t) \wedge \lim_{i \rightarrow \infty} \{(B^{(i)})^\circ \otimes U(0, t)\} + \Sigma(t) \\ &= A(t) \wedge \lim_{i \rightarrow \infty} (B^{(i)})^\circ \otimes U(0, t) + \Sigma(t) \\ &= A(t) \wedge (\lim_{i \rightarrow \infty} B^{(i)})^\circ \otimes U(0, t) + \Sigma(t) \end{aligned}$$

We have used Lemma 15 in the fourth line. The last equality is only possible if changing order of the limit with the $(\)^\circ$ -operator is possible. As this is in general not true we give a proof for that:

Assume an interval $[s, r]$ on which A is continuous and fix some $t \in [s, r]$. We know:

$$B^{(i)}(t) = A(t) \wedge A \otimes U(0, t) + \Sigma(t) \wedge \dots \wedge A \otimes U^{(i)}(0, t) + i\Sigma(t)$$

As the min-plus convolution and the minima preserve continuity, we have that $B^{(i)}$ is continuous in t as well. Assume now there would be a (unbounded) sequence i_k , such that:

$$\lim_{i \rightarrow \infty} B^{(i)}(t) = \lim_{k \rightarrow \infty} A \otimes U^{(i_k)}(0, t) + i_k \Sigma(t)$$

This could be interpret, that we always need to consider new terms in the sequence of minima to calculate $\lim_{i \rightarrow \infty} B^{(i)}(t)$. It would hold then

$$A \otimes U^{(i_k)}(0, t) + i_k \Sigma(t) = B^{(i_k)}(t) \leq A(t),$$

following:

$$i_k \Sigma(t) \leq \sup_{0 \leq s \leq t} \{A(s, t) - U^{(i_k)}(s, t)\} \quad \forall i \in \mathbb{N}_0.$$

By the non-negativity of U we would have eventually:

$$\Sigma(t) \leq \sup_{0 \leq s \leq t} \left\{ \frac{A(s, t)}{i_k} - \frac{U^{(i_k)}(s, t)}{i_k} \right\} \xrightarrow{k \rightarrow \infty} 0$$

meaning $\Sigma(t) \leq 0$, which contradicts the assumption of $\Sigma(t) \geq \Sigma_{min} > 0$. Hence an unbounded sequence i_k as above cannot exist and there exists always some index i^* with $\lim_{i \rightarrow \infty} B^{(i)}(t) = B^{(i^*)}(t)$. This allows to change the limit with the $(\)^\circ$ -operator in all times, in which A is continuous.

¹It is easy to see, that, if $B^{(i)}(t) \geq B^{(i-1)}(t)$ for all t , then also $(B^{(i)})^\circ(t) \geq (B^{(i-1)})^\circ(t)$ for all t .

Next consider by \mathcal{A} the times at which A is discontinuous. As

$$B^{(k)}(t) = A(t) \wedge A \otimes U(0, t) + \Sigma(t) \wedge \dots \wedge A \otimes U^{(k)}(t) + k\Sigma(t)$$

and U is proper, the times at which $B^{(k)}$ is discontinuous form a subset of \mathcal{A} , we call it \mathcal{A}_k and $\mathcal{A}_\infty := \bigcap_k \mathcal{A}_k$. On $t \in \mathcal{A}_\infty$ we have (since every part of the above representation is continuous, except for A):

$$B^{(k)}(s) = A(s)$$

for all $s \in [t - \varepsilon, t)$ and some $\varepsilon > 0$. Hence, if $\lim_{k \rightarrow \infty} B^{(k)}(t)$ is discontinuous on some t^* , we have:

$$\lim_{k \rightarrow \infty} \lim_{t \nearrow t^*} B^{(k)}(t) = \lim_{k \rightarrow \infty} \lim_{t \nearrow t^*} A(t) = A^\circ(t) = \lim_{t \nearrow t^*} \lim_{k \rightarrow \infty} B^{(k)}(t)$$

□

The next theorem derives the throttle's service description:

THEOREM 14. *Under the above conditions the throttle is a dynamic U_\wedge -server with*

$$(4.0.12) \quad U_\wedge(s, t) := \bigwedge_{i=0}^{\infty} (U')^{(i)}(s, t),$$

and $U'(s, t) = U(s, t) + \Sigma^\circ(t)$. I.e., for all B fulfilling (4.0.10) holds $B(t) \geq A^\circ \otimes U_\wedge(0, t)$.

Eventually, we find that the results for the bivariate, continuous time scenario, parallel those for discrete time and only differ in a modified service description.

The proof of Theorem 14 needs a Lemma about continuity:

LEMMA 15. *Let A_k be a decreasing sequence of flows, i.e. $A_k(s) \geq A_{k+1}(s)$ for all $s \in \mathbb{R}_0^+$ and denote $\lim_{k \rightarrow \infty} A_k(s) = A'(s)$. Further let U be some non-negative service. Then:*

$$\lim_{k \rightarrow \infty} (A_k \otimes U)(s, t) = A' \otimes U(s, t) \quad \forall s \leq t \in \mathbb{R}_0^+$$

Further, if $U_k(s, t)$ is a decreasing sequence of services ($U_k(s, t) \geq U_{k+1}(s, t)$ for all $s \leq t \in \mathbb{R}_0^+$), we have for each flow A :

$$\lim_{k \rightarrow \infty} (A \otimes U_k)(s, t) = A \otimes U'(s, t) \quad \forall s \leq t \in \mathbb{R}_0^+$$

PROOF. We start with the first assertion, the second follows in the same manner. Let U be some arbitrary service. We show first the existence of the left-handed-side limit. Let $s \leq t$ be arbitrary and r^* the index minimizing $\inf_{s \leq r \leq t} \{A_k(s, r) + U(r, t)\}$, then:

$$\begin{aligned} & \inf_{s \leq r \leq t} \{A_k(s, r) + U(r, t)\} - \inf_{s \leq r \leq t} \{A_{k+1}(s, r) + U(r, t)\} \\ & \geq A_k(s, r^*) + U(r^*, t) - A_{k+1}(s, r^*) - U(r^*, t) \\ & \geq A_k(r^*) - A_{k+1}(r^*) \\ & \geq 0 \end{aligned}$$

and hence the sequence $(A_k \otimes U)_{k \in \mathbb{N}_0}$ is decreasing. Further we have that $A_k \otimes U \geq 0$ for all k , ensuring the existence of $\lim_{k \rightarrow \infty} A_k \otimes U$. We show the equality in two steps:

“ \leq ”: Let r^* be the index minimizing $\inf_{s \leq r \leq t} \{A'(s, r) + U(r, t)\}$. Then:

$$\begin{aligned} \lim_{k \rightarrow \infty} \inf_{s \leq r \leq t} \{A_k(s, r) + U(r, t)\} &\leq \lim_{k \rightarrow \infty} A_k(s, r^*) + U(r^*, t) \\ &= A'(s, r^*) + U(r^*, t) \\ &= \inf_{s \leq r \leq t} \{A'(s, r) + U(r, t)\} \end{aligned}$$

“ \geq ”: From the monotonicity we have:

$$\begin{aligned} \lim_{k \rightarrow \infty} \inf_{s \leq r \leq t} \{A_k(s, r) + U(r, t)\} &\geq \lim_{k \rightarrow \infty} \inf_{s \leq r \leq t} \{A'(r) + U(r, t)\} - A_k(s) \\ &= \inf_{s \leq r \leq t} \{A'(s, r) + U(r, t)\} \end{aligned}$$

□

Now we can give the proof of Theorem 14

PROOF. Let B be some flow fulfilling (4.0.10). We know from the previous theorem that $B \geq B'$ holds. Hence it is sufficient to prove $B'(t) \geq A^\circ \otimes U_{\mathcal{T}}(0, t)$. Denote $U_{\mathcal{T}_k}(s, t) = \bigwedge_{i=0}^k (U + \tilde{\Sigma})^{(i)}(s, t)$. Assume we would have $B^{(k)} \geq A^\circ \otimes U_{\mathcal{T}_k}$ for all $k \in \mathbb{N}_0$, then

$$B' = \lim_{k \rightarrow \infty} B^{(k)} \geq \lim_{k \rightarrow \infty} A^\circ \otimes U_{\mathcal{T}_k} = A^\circ \otimes U_{\mathcal{T}}$$

where we have used the second part of Lemma 15. So, we only have to show the above assumption. For $k = 0$ we have

$$B^{(0)} = A \geq A^\circ \otimes \mathbf{1} = A \otimes \bigwedge_{i=0}^0 (U + \tilde{\Sigma})^{(0)}$$

Given $B^{(k)} \geq A^\circ \otimes U_{\mathcal{T}_k}$ for some $k \in \mathbb{N}_0$ we get for all $t > 0$:

$$\begin{aligned} B^{(k+1)}(t) &= A(t) \wedge (B^{(k)})^\circ \otimes U(0, t) + \Sigma(t) \\ &\geq A^\circ(t) \wedge (A^\circ \otimes U_{\mathcal{T}_k}) \otimes U(0, t) + \Sigma(t) \\ &\geq A^\circ \otimes \mathbf{1}(0, t) \wedge A^\circ \otimes (U_{\mathcal{T}_k} \otimes (U + \tilde{\Sigma}))(0, t) \\ &= A^\circ \otimes (\mathbf{1}(0, t) \wedge U_{\mathcal{T}_k} \otimes (U + \tilde{\Sigma}))(0, t) \\ &= A^\circ \otimes U_{\mathcal{T}_{k+1}}(0, t) \end{aligned}$$

In the second line we used

$$\begin{aligned} (B^{(k)})^\circ \otimes U(0, t) &= \inf_{0 \leq s \leq t} \{(B^{(k)})^\circ(s) + U(s, t)\} \\ &\geq 0 + U(0, t) \wedge \inf_{0 < s \leq t} \{A^\circ \otimes U_{\mathcal{T}_k}(0, s) + U(s, t)\} \\ &= \inf_{0 \leq s \leq t} \{A^\circ \otimes U_{\mathcal{T}_k}(0, s) + U(s, t)\}, \end{aligned}$$

since $A^\circ \otimes U_{\mathcal{T}_k}(0, 0) = 0$.

In the third line we used:

$$\begin{aligned}
& (A^\circ \otimes U_{\mathcal{T}_k}) \otimes U(0, t) + \Sigma(t) \\
&= A^\circ \otimes (U_{\mathcal{T}_k} \otimes U)(0, t) + \tilde{\Sigma}(0, t) \\
&= \inf_{0 \leq s \leq t} \{A^\circ(s) + \inf_{s \leq r \leq t} \{U_{\mathcal{T}_k}(s, r) + U(r, t)\}\} + \tilde{\Sigma}(0, t) \\
&\geq \inf_{0 \leq s \leq t} \{A^\circ(s) + \inf_{s \leq r \leq t} \{U_{\mathcal{T}_k}(s, r) + U(r, t) + \tilde{\Sigma}(r, t)\}\} \\
&= A^\circ \otimes (U_{\mathcal{T}_k} \otimes (U + \tilde{\Sigma}))(0, t)
\end{aligned}$$

This works, since $\Sigma^\circ(r) \geq 0$ for all $r \geq 0$.

□

CHAPTER 5

Subadditive Service

In this chapter, we investigate subadditive service descriptions in WFC systems. An example for a subadditive server is given by a constant rate server $U_o(s, t) = r_U(t - s)$ with some positive rate r_U serving a crossflow A_U . As discussed above, we need to replace the crossflow A_U by \tilde{A}_U if we use continuous time. Because of this and the somewhat different results in continuous time, we distinguish between both time models in this discussion.

5.1. Discrete Time

Assume we have a subadditive service description U_{sub} covering the tandem of the service element U and the placeholder in Figure 2.0.1, i.e., $U_{fb}(s, t) = U_{sub}(s, t) + \Sigma(t)$. We then have

$$\begin{aligned} & U_{fb} \otimes U_{fb}(s, t) \\ &= \min_{s \leq r \leq t} \{U_{sub}(s, r) + \Sigma(r) + U_{sub}(r, t) + \Sigma(t)\} \\ &\geq U_{fb}(s, t). \end{aligned}$$

Applying this in Theorem 5 the throttle can be described by:

$$U_{\wedge}(s, t) = \mathbf{1}(s, t) \wedge U_{fb}(s, t),$$

and with Theorem 3 we eventually have for the whole system the service description:

$$\begin{aligned} U_{sys}(s, t) &= U_{\wedge} \otimes U(s, t) \\ &\geq U(s, t) \wedge U_{sub} \otimes U(s, t) + \min_{s \leq r \leq t} \Sigma(r) \\ &\geq U_{sub} \otimes U(s, t), \end{aligned}$$

where we use the monotonicity of the min-plus convolution, i.e., it holds $U \otimes V(s, t) \leq U(s, t)$ for all V with $V(t, t) = 0$ for all t .

Particularly, we have for $U_{sub} = U$ the system description:

$$U_{sys}(s, t) \geq U(s, t),$$

i.e., the system can be analysed as the unthrottled one and Theorem 8 gives end-to-end delay bounds.

5.2. Continuous Time

To make a long story short: in principle, the results of discrete time carry over to the setting of continuous time; however, the slightly different characterization of

leftover service (Lemma 12) interferes with subadditivity. One has in contrast to 10 for a subadditive service U_o , crossflow A_U and arbitrary times $s \leq r \leq t$:

$$\begin{aligned} U(s, r) + U(r, t) &= U_o(s, r) - \tilde{A}_U(s, r) + U_o(r, t) - \tilde{A}_U(r, t) \\ &\geq U_o(s, t) - \tilde{A}_U(s, t) - \tilde{A}_U(r, r) \\ &= U(s, t) - \tilde{A}_U(r, r) \end{aligned}$$

Hence, U remains subadditive if $\tilde{A}_U(r, r) = 0$. This motivates the following definition.

DEFINITION 16. A dynamic U -server is *subadditive with defect A* (abbreviated *A -subadditive*), if for all $s \leq r \leq t$

$$U(s, r) + U(r, t) \geq U(s, t) - \tilde{A}(r, r).$$

The intuition here is, that U is “mostly” subadditive, but the subadditivity can be violated, if the crossflow A_U contains an instantaneous burst. Although the above definition is weaker than conventional subadditivity, it still lends itself to an analysis of WFC systems. The key here is that the window process can compensate for the subadditivity’s defect.

THEOREM 17. *Let U be an A_U -subadditive dynamic server. For the k -th ($k > 0$) self-convolution of $U'(s, t) := U(s, t) + \Sigma^\circ(t)$ holds:*

$$(5.2.1) \quad (U')^{(k)}(s, t) \geq U'(s, t) + (k - 1) \inf_{s \leq r \leq t} \{\Sigma^\circ(r) - \tilde{A}_U(r, r)\}$$

PROOF. We prove this by induction over k . The statement is obviously true for $k = 1$. Assuming (5.2.1) holds for k we have for $k + 1$:

$$\begin{aligned} (U')^{(k+1)}(s, t) &= \inf_{s \leq r \leq t} \{(U')^{(k)}(s, r) + U'(r, t)\} \\ &\geq \inf_{s \leq r \leq t} \{U(s, r) + \Sigma^\circ(r) + U(r, t) + \Sigma^\circ(t) \\ &\quad + (k - 1) \inf_{s \leq q \leq r} \{\Sigma^\circ(q) - \tilde{A}_U(q, q)\}\} \\ &\geq \inf_{s \leq r \leq t} \{U'(s, t) + \Sigma^\circ(r) - \tilde{A}_U(r, r) \\ &\quad + (k - 1) \inf_{s \leq q \leq r} \{\Sigma^\circ(q) - \tilde{A}_U(q, q)\}\} \\ &\geq U'(s, t) + \inf_{s \leq r \leq t} \{\Sigma^\circ(r) - \tilde{A}_U(r, r)\} \\ &\quad + (k - 1) \inf_{s \leq r \leq t} \inf_{s \leq q \leq r} \{\Sigma^\circ(q) - \tilde{A}_U(q, q)\} \\ &= U'(s, t) + k \inf_{s \leq r \leq t} \{\Sigma^\circ(r) - \tilde{A}_U(r, r)\}, \end{aligned}$$

which proves the theorem. \square

Similar to the discrete time setting we now assume an A_U -subadditive service description U_{sub} covering the service element U and the placeholder in Figure 2.0.1

such that $U_{fb}(s, t) = U_{sub}(s, t) + \Sigma(t)$. Applying Theorem 17 in (4.0.12) we obtain:

$$\begin{aligned} U_{\wedge}(s, t) &= \bigwedge_{k=0}^{\infty} U_{fb}^{(k)}(s, t) \\ &\geq \mathbf{1}(s, t) \wedge \bigwedge_{k=1}^{\infty} U_{sub}(s, t) + \Sigma(t) \\ &\quad + (k-1) \inf_{s \leq r \leq t} \{\Sigma^{\circ}(r) - \tilde{A}_U(r, r)\}, \end{aligned}$$

which in turn, depending on the sign of the involved infimum, reduces to two values only

$$(5.2.2) \quad U_{\wedge}(s, t) \geq \begin{cases} 0 & \text{if } \exists r \in [s, t] \text{ with } \tilde{A}_U(r, r) > \Sigma^{\circ}(r) \\ \mathbf{1}(s, t) \wedge U_{sub}(s, t) + \Sigma(t) & \text{else} \end{cases}$$

resulting in

$$U_{sys}(s, t) \geq \begin{cases} 0 & \text{if } \exists r \in [s, t] \text{ with } \tilde{A}_U(r, r) > \Sigma^{\circ}(r) \\ U_{sub}(s, t) & \text{else} \end{cases}$$

Hence, if we know that the maximal burst of A_U in some interval does not exceed the current window size, we can analyse the system as an unthrottled one. Calculating that probability is thus key when we want to calculate probabilistic performance bounds as in Theorem 8. This is formulated in the following theorem.

THEOREM 18. *Assume $U_{fb}(s, t) = U_{sub}(s, t) + \Sigma(t)$ and the conditions of Theorem 8 hold for A and U_{sub} . Further, let U_{sub} be A_U -subadditive. The following end-to-end delay bound holds for any $t, T \geq 0$:*

$$\begin{aligned} &\mathbb{P}(d(t) > T) \\ &\leq e^{\theta(\sigma_U(\theta) + \sigma_A(\theta) + \rho_U(\theta)(T - \delta))} (1 - e^{\delta\theta\rho(\theta)})^{-1} + \mathbb{P}(E), \end{aligned}$$

where $\delta > 0$ is a free discretization parameter and

$$\mathbb{P}(E) := \mathbb{P}\left(\sup_{r \in [0, t+T]} \{\tilde{A}_U(r, r) - \Sigma(r)\} > 0\right).$$

PROOF. We start by applying the law of total probability, conditioning on the event E :

$$(5.2.3) \quad \begin{aligned} &\mathbb{P}(d(t) > T) \\ &\leq \mathbb{P}(d(t) > T \mid E)\mathbb{P}(E) + \mathbb{P}(d(t) > T \mid \neg E)\mathbb{P}(\neg E) \\ &\leq \mathbb{P}(E) + \mathbb{P}(d(t) > T \mid \neg E) \end{aligned}$$

As we observed for discrete time the expression $\mathbb{P}(d(t) > T \mid \neg E)$ can be treated by applying Theorem 8 on U_{sub} . The version of this result in continuous time requires

a discretization parameter δ and reads:

$$\begin{aligned}
& \mathbb{P}(d(t) > T \mid \neg E) \\
& \leq \mathbb{P}\left(\sup_{0 \leq s \leq t+T} \{A(r, t) - U_{sub}(r, t+T)\} > 0\right) \\
& \leq \mathbb{P}\left(\max_{0 \leq i \leq N-1} \{A(i\delta, t) - U_{sub}((i+1)\delta, t+T)\} > 0\right) \\
& \leq \sum_{i=0}^{N-1} \mathbb{E}(e^{\theta A(i\delta, t)}) \mathbb{E}(e^{-\theta U_{sub}((i+1)\delta, t+T)}) \\
& \leq \sum_{i=0}^{N-1} e^{\theta \rho_A(\theta) \delta(N-i) + \theta \sigma_A(\theta)} e^{\theta \rho_U(\theta)(T-\delta) + \theta \rho_U(\theta) \delta(N-i) + \theta \sigma_U(\theta)} \\
& = e^{\theta(\sigma_A(\theta) + \sigma_U(\theta) + \rho_U(\theta)(T-\delta))} \sum_{j=1}^N e^{\theta \rho_U(\theta) \delta j} \\
& \leq e^{\theta(\sigma_A(\theta) + \sigma_U(\theta) + \rho_U(\theta)(T-\delta))} \cdot (1 - e^{\delta \theta \rho_U(\theta)})^{-1}.
\end{aligned}$$

Here, we divided the interval $[0, t]$ into N slots each of length δ . \square

To obtain a numerical value for the above delay bound, we need to find an appropriate bound for the event E . This depends on the type of arrivals A :

EXAMPLE 19. Assume the jumps occur at integer times $n = 1, 2, \dots$ and their jump sizes have an i.i.d. distribution F_J . Further define $\Sigma_{min}(t) = \inf_{0 \leq s \leq t} \Sigma(s)$. Then:

$$\mathbb{P}\left(\max_{1 \leq n \leq \lfloor t+T \rfloor} \tilde{A}_U(n, n) > \Sigma_{min}(t)\right) = 1 - F_J(\Sigma_{min}(t))^{\lfloor t+T \rfloor}.$$

We can extend this example to bursts at random times: assume the timing of bursts follows a Poisson process. Splitting by the number of bursts we have:

$$\mathbb{P}(E) = \sum_{k=0}^{\infty} (1 - F_J(\Sigma_{min}(t))^{\lfloor k \rfloor}) \mathbb{P}([0, t+T] \cap J = k).$$

In this section, we showed that subadditive service is sufficient to make the analysis of WFC systems in MGF-calculus possible. Subadditivity reduces the complexity of describing \bar{U}_{fb} immensely, such that a closed-form description of the end-to-end service is possible. On the other hand, it is a strong assumption to have a subadditive service description for U_{sub} in $U_{fb} = U_{sub} + \Sigma$. We explore the scope of this approach in the next section.

5.3. Comments on Network Analysis

Excluding the window element we can consider the feedback loop as any kind of network, in which B enters and C' departs (see Figure 2.0.1). The goal is to find a subadditive service description of this network. We can intuitively extend the notion of A -subadditivity to networks, if they can be represented by some A -subadditive service

$$C(t) \geq B \otimes U_{e2e}(0, t) \quad \forall t \geq 0.$$

To answer which kind of networks are subadditive we check two network operations with respect to preserving subadditivity. The first considers demultiplexing of flows and the second is the concatenation of service elements.

THEOREM 20. *Consider a subadditive dynamic U -server and a flow A . The leftover service $U_l = U - \tilde{A}$ is A -subadditive. If U is A_U -subadditive, the leftover service is $A + A_U$ -subadditive.*

PROOF. We prove the second part only, as the first part can be considered as a special case with $A_U = 0$. Let $s \leq r \leq t$ be arbitrary:

$$\begin{aligned} U_l(s, r) + U_l(r, t) &= U(s, r) + U(r, t) - \tilde{A}(s, r) - \tilde{A}(r, t) \\ &\geq U(s, t) - \tilde{A}_U(r, r) - \tilde{A}(r, r) - \tilde{A}(s, t) \\ &= U_l(s, t) - (\tilde{A}_U(r, r) + \tilde{A}(r, r)) \end{aligned}$$

□

Next, we give a subadditive description for concatenated service elements. However, we first need a definition, which is slightly stricter than subadditivity:

DEFINITION 21. A dynamic U -server is called *separable*, if there exists a subadditive U_o and a flow A_U , such that

$$U(s, t) = U_o(s, t) - \tilde{A}_U(s, t).$$

Note that a separable dynamic U -server is always A_U -subadditive, but the converse does not hold in general. Often, we are able to represent U_o and A_U of a separable server in an additive form:

$$(5.3.1) \quad U_o(s, t) = \int_s^t u_o(x) dx$$

$$(5.3.2) \quad A_U(s, t) = \int_s^t a_U(r) dr + \sum_{r \in [s, t]} \tilde{A}_U(r, r),$$

for some integrable functions u_o and a_U .

We assume such representations for the rest of this section. This enables the following theorem.

THEOREM 22. *Assume separable dynamic U - and V - servers. In tandem they form a separable dynamic W -server with*

$$\begin{aligned} W_o(s, t) &:= \sum_{p=s}^{t-1} U_o(p, p+1) \wedge V_o(p, p+1) \\ A_W(s, t) &:= \sum_{p=s}^{t-1} A_U(p, p+1) \vee A_V(p, p+1). \end{aligned}$$

in discrete time and

$$\begin{aligned} W_o(s, t) &:= \int_s^t u_o(x) \wedge v_o(x) dx \\ \tilde{A}_W(s, t) &:= \int_s^t a_U(x) \vee a_V(x) dx + \sum_{q \in p[s, t]} \tilde{A}_U(q, q) + \tilde{A}_V(q, q) \end{aligned}$$

for continuous time. W is $A_U + A_V$ -subadditive.

PROOF. We give the proof for continuous time as the one for discrete time is just a special case. We know that the tandem is a $U \otimes V$ -server, further we have for all $s \leq t$:

$$\begin{aligned}
& U \otimes V(s, t) \\
& \geq \inf_{s \leq r \leq t} \{U_o(s, r) + V_o(r, t)\} - \sup_{s \leq r \leq t} \{\tilde{A}_U(s, r) + \tilde{A}_V(r, t)\} \\
& \geq \inf_{s \leq r \leq t} \{W_o(s, r) + W_o(r, t)\} - \sup_{s \leq r \leq t} \left\{ \int_s^t a_U(x) \vee a_V(x) dx \right. \\
& \quad \left. + \sum_{q \in [s, r]} \tilde{A}_U(q, q) + \sum_{q \in [r, t]} \tilde{A}_V(q, q) \right\} \\
& \geq W_o(s, t) - \tilde{A}_W(s, t),
\end{aligned}$$

and hence W is a separable dynamic server for the tandem as well. It is easy to see that its defect equals $A_U + A_V$. \square

The above two operations allow to deal with a class of subadditive networks.

LEMMA 23. *Any feedforward network consisting of separable dynamic U_i -servers ($i \in 1, \dots, N$) with strict priority scheduling and flows A_j ($j \in 1, \dots, M$) is subadditive with appropriate defect.*

PROOF. W.l.o.g we assume all $A_{U_i} = 0$. We prove only the more general continuous case. Pick any flow A_j and denote the indices of the service elements on its path by i_1, i_2, \dots, i_K . Assume for a moment only one cross-flow A_l interfering with A_j , i.e., it has higher priority on at least one of the service elements. We show now, that extending A_l and increasing its priority above A_j on the whole path preserves the bounds' correctness. Denote by U_k the available service for flow A_j at service element i_k , before subtracting A_l . Denote by B_j the departures of A_j from the network, then we have

$$B_j \geq A_j \otimes (U_1 - A_{l,1} \otimes \dots \otimes U_K - A_{l,K}),$$

where some of the crossflows $A_{l,k}$ are equal to zero, if the crossflow does not interfere with the service for A_j at service element U_k . Let k' be the index of the service element at which A_l interferes with the path of A_j for the first time. Due to causality we must have $A_{l,k'} \geq A_{l,k}$ for any $k \in \{1, \dots, K\}$. Hence:

$$B_j \geq A_j \otimes (U_1 - A_{l,k'} \otimes \dots \otimes U_K - A_{l,k'}),$$

and Theorem 22 leads to a single separable service by:

$$\begin{aligned}
W_o(s, t) &= \int_s^t \bigwedge_{k=1}^K u_{i_k}(x) dx \\
\tilde{A}_W(s, t) &= \int_s^t a_{l,k'}(x) dx + K \sum_{q \in [s, t]} \tilde{A}_{l,k'}(q, q),
\end{aligned}$$

and W is subadditive with defect $KA_{l,k'}$. We can now perform this procedure again, but for all flows interfering with A_j leading to:

$$B_j \geq A_j \otimes (U_1 - \sum_{l \in \mathcal{L}} A_{l,k'_l}, \dots, U_K - \sum_{l \in \mathcal{L}} A_{l,k'_l})$$

where \mathcal{L} is the set of indices of all interfering flows. By applying Theorem 22 as before, we obtain a single separable service element W with the same W_o and

$$\tilde{A}_W(s, t) = \int_s^t \sum_{l \in \mathcal{L}} a_{l, k'_l}(x) dx + K \sum_{q \in [s, t]} \sum_{l \in \mathcal{L}} \tilde{A}_{l, k'_l}(q, q)$$

The defect on its subadditivity is $K \sum_{l \in \mathcal{L}} A_{l, k'_l}$. \square

In the above proof, we used that we can extend any cross-flow interfering with the flow of interest on the latter's whole path. Depending on the actual network topology, this method can become arbitrarily loose. We know from DNC that, for example in nested topologies, better service descriptions are available [16]. Such improvements rely on the structure of the considered network and a suitable ordering demultiplexing and convolution operations. These service descriptions, however, are in general not subadditive.

While the above lemma characterizes a fairly large class of networks to be subadditive, the question rises, if one can improve Lemma 23 similarly, without losing subadditivity. We provide a simple example for this, if the crossflows can be partitioned. To that end, consider again the service elements i_k lying on the path of our flow of interest. We denote all flows sharing the same service element on their way through the network by \mathcal{F}_k . We say the cross-flows can be partitioned, if we can find a (non-trivial) partition $\bigcup P_m = \{i_1, \dots, i_K\}$, such that

$$\bigcup_{i_k \in P_{m_1}} \mathcal{F}_k \cap \bigcup_{i_k \in P_{m_2}} \mathcal{F}_k = \emptyset$$

We denote all flows lying in one set of this partition by $\mathcal{P}_m := \bigcup_{i_k \in P_m} \mathcal{F}_k$ and the sum of the corresponding flows by $\mathcal{A}_m(s, t) = \sum_{j \in \mathcal{P}_m} A_j(s, t)$. We can now improve on the above lemma, without losing subadditivity:

COROLLARY 24. *Consider a flow of interest A_j and its corresponding partitioned network. The subadditive dynamic W -server for flow A_j can be improved to:*

$$W_o(s, t) = \int_s^t \bigwedge_{k=1}^K u_{i_k}(x) dx$$

$$\tilde{A}_W(s, t) = \int_s^t \bigvee_m \mathcal{A}_m(x) dx + \sum_m |P_m| \sum_{q \in J \cap [s, t]} \tilde{\mathcal{A}}_m(q, q)$$

with defect $\sum_m |P_m| \mathcal{A}_m$.

PROOF. Assume first that we can order the partition, such that for all $i_k \in P_m$ holds either $i_k < i_{k'}$ or $i_k > i_{k'}$ for all $i_{k'} \in P_{m'}$ and $m \neq m'$. We use the partition on the involved service elements and apply Lemma 23 on each set P_m locally, to achieve separable service descriptions W_m with:

$$W_{o,m}(s, t) = \int_s^t \bigwedge_{k \in P_m} u_k(x) dx$$

$$\tilde{A}_{W,m}(s, t) = \int_s^t \mathcal{A}_m(x) dx + |P_m| \sum_{q \in J \cap [s, t]} \tilde{\mathcal{A}}_m(q, q)$$

and defect $|P_m|A_m$. Then, in a second step, we apply Lemma 23 again on the service elements W_m . If the partitions cannot be ordered as above, it is easy to see, that we can find a finer partition, which can be ordered, and delivers the same result. \square

To sum up, we have proven the existence of subadditive service descriptions for any feed-forward network under strict priority scheduling. Those in turn can be plugged into the feedback-inequality and allow us to analyse the end-to-end WFC service. Corollary 24 gives a simple example on how one can take advantage of the network's topology to achieve better subadditive service descriptions. More sophisticated exploitation of topological characteristics are left to future work.

5.4. Numerical Evaluation

We consider the following scenario: The feedback loop consists of two service elements. The service element U is a constant rate server with rate u , which serves a high-priority cross-flow A_U . The service element V is a constant rate server with rate v and cross-flow A_V . We define a certain kind of arrival process in this section and use continuous time, to display our results in their generality. All arrival flows in our scenario are defined by the compound of two processes each. We illustrate this for the arrival flow A : we use two sequences of i.i.d. exponentially distributed random variables $(X_n)_{n \in \mathbb{N}}$ and $(X_n^j)_{n \in \mathbb{N}}$. Now define A by $A(t) = A^c(t) + A^j(t)$, such that A^c is a piecewise linear function with rates X_n and A^j is a sequence of jumps at times \mathbb{N}_0 of size X_n^j , i.e.:

$$A^c(t) - A^c(s) = (t - s)X_n, \quad s, t \in [n, n + 1] \text{ and } n \in \mathbb{N}_0$$

$$\tilde{A}^j(s, t) = \begin{cases} X_n^j & \text{if } s = t \text{ and } s \in \mathbb{N}_0 \\ 0 & \text{if } t - s < 1 \text{ and } s \notin \mathbb{N}_0 \end{cases}$$

We assume in this section that all sequences X_n, X_n^j, \dots are stochastically independent. To obtain an MGF-bound for these arrivals we need

LEMMA 25. *Be A as defined above with X_n and X_n^j being any sequence of i.i.d. random variables, with $\mathbb{E}(e^{\theta X_n}) \leq e^{\theta \rho_c(\theta)}$ and $\mathbb{E}(e^{\theta X_n^j}) \leq e^{\theta \rho_j(\theta)}$, respectively, for all $0 < \theta < \theta^*$, some $\theta^* > 0$ and ρ being an increasing function. Then we have for all $0 \leq s \leq t$ and $0 < \theta < \theta^*$:*

$$\mathbb{E}(e^{\theta A(s,t)}) \leq e^{\theta(t-s)(\rho_c(\theta) + \rho_j(\theta)) + \theta \rho_j(\theta)}$$

PROOF. The proof is just keeping track of the single MGFs of all involved X_n and X_n^j , if $s, t \notin \mathbb{N}_0$:

$$\begin{aligned}
\Phi_{A(s,t)}(\theta) &= \mathbb{E}(e^{\theta((\lceil s \rceil - s)X_{\lceil s \rceil} + \sum_{n=\lceil s \rceil}^{\lfloor t \rfloor} (X_n + X_n^j) + (t - \lfloor t \rfloor)X_{\lceil t \rceil})}) \\
&= \mathbb{E}(e^{\theta((\lceil s \rceil - s)X_{\lceil s \rceil})}) \mathbb{E}(e^{\theta(\sum_{n=\lceil s \rceil}^{\lfloor t \rfloor} X_n + X_n^j)}) \mathbb{E}(e^{\theta((t - \lfloor t \rfloor)X_{\lceil t \rceil})}) \\
&\leq e^{\theta(\lceil s \rceil - s)\rho_c(\theta(\lceil s \rceil - s))} \cdot \prod_{n=\lceil s \rceil}^{\lfloor t \rfloor} \mathbb{E}(e^{\theta X_n + X_n^j}) \cdot e^{\theta(t - \lfloor t \rfloor)\rho_c(\theta(t - \lfloor t \rfloor))} \\
&\leq e^{\theta(\lceil s \rceil - s)\rho_c(\theta)} \prod_{n=\lceil s \rceil}^{\lfloor t \rfloor} e^{\theta\rho_c(\theta) + \theta\rho_j(\theta)} e^{\theta(t - \lfloor t \rfloor)\rho_c(\theta)} \\
&= e^{\theta(t-s)(\rho_c(\theta) + \rho_j(\theta))}
\end{aligned}$$

However if $s, t \in \mathbb{N}_0$ we need to replace $\prod_{n=\lceil s \rceil}^{\lfloor t \rfloor} \mathbb{E}(e^{\theta X_n^j})$ by $\prod_{n=s}^t \mathbb{E}(e^{\theta X_n^j}) \leq e^{\theta\rho_j(\theta)(t-s+1)}$. \square

If not specified differently we use the following parameters in this section: the exponential rates of the processes X_n, X_n^j, \dots are all chosen equally to $\lambda = 8$ and the service rates are $u = v = 1$. We evaluate the bound at time $t = 5$ (note: this is only important for the throttled case, as the unthrottled bound is independent of t) and ask for a delay of $T = 10$. We numerically optimized over the parameter θ and have chosen a reasonable discretization parameter δ by hand. Checking the bound derived in Equation (5.2.3), we see that the difference between the throttled and unthrottled case comes in form of adding the probability $\mathbb{P}(E)$. Assuming a constant window size $\Sigma(t) \geq \Sigma$, this is given in our case by:

$$1 - (F_{2X^j}(\Sigma))^{\lfloor t+T \rfloor} \leq 1 - (1 - e^{-\lambda\Sigma} - e^{-\lambda\Sigma}\lambda\Sigma)^{(t+T)}$$

The delay bound for the unthrottled case is calculated from Theorem 8.

5.4.1. Convergence to Unthrottled System. When we compare the bounds for the throttled and the unthrottled system we see two differences. The first is the additive component $\mathbb{P}(E)$, which decreases for increasing window sizes Σ . The second is a difference in the delay-calculation: the throttled system uses the service resulting from Theorem 22, which takes the server V into account. Figure 5.4.1 shows in black lines, how the bounds evolve for an increasing window size and different rates $u = 1, 1.2, 1.5$. One sees that the throttled systems converge to a bound which lies above the bound of the corresponding unthrottled systems (red lines), which is due to adding V to the system.

5.4.2. Dependence on Delay. Next, we investigate how the delay parameter T influences the scenario, as it appears in both parts of the bound for the throttled system (improving the delay-bound for increasing T , while worsening the bound in the probability $\mathbb{P}(E)$). To make this effect visible, we keep the window size at $\Sigma = 5$ and the other parameters as before. In Figure 5.4.2, one can observe how increasing the parameter T first leads to an improvement of the bound; yet, increasing the delay, the throttled part starts to dominate the bound and it slowly worsens with further increase of T .

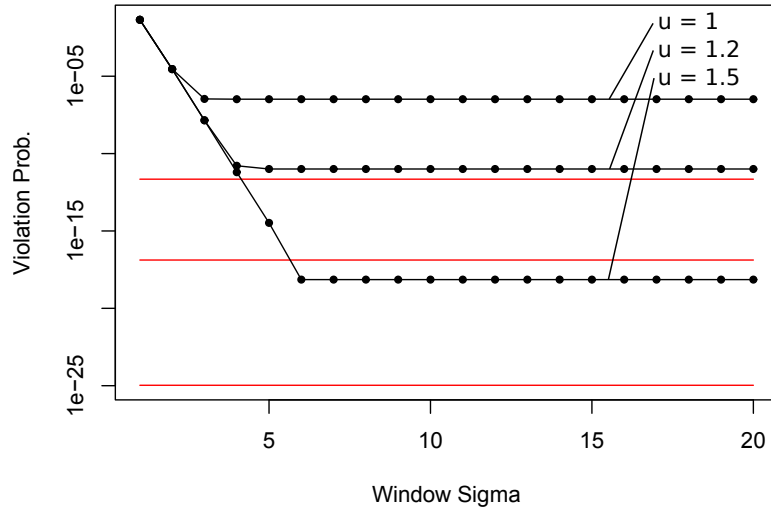


FIGURE 5.4.1. Convergence of throttled systems for increasing window-sizes. The red lines indicate the corresponding unthrottled systems.

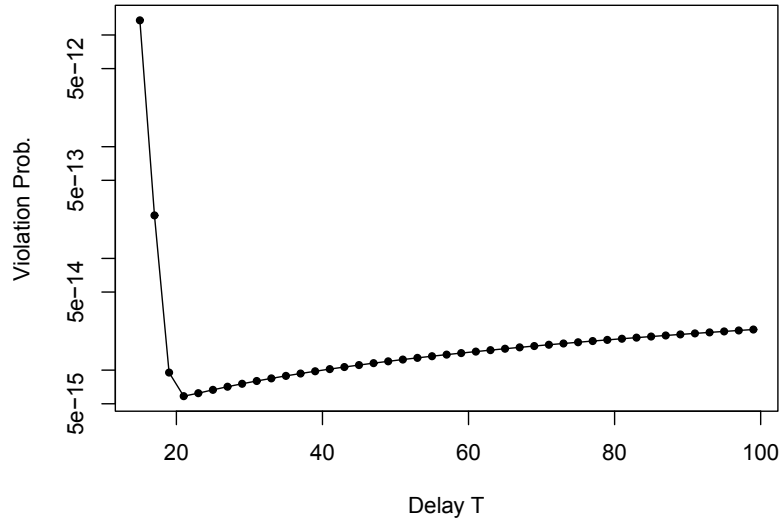


FIGURE 5.4.2. Influence of Delay T on the bound.

CHAPTER 6

General Case

We revisit the solution of Theorem 5. Without further assumptions, we need to find an MGF-bound in the sense of (2.0.3) for the end-to-end service:

$$U_{sys} = \left(\bigwedge_{n=0}^{\infty} (U_{fb})^{(n)} \right) \otimes U(s, t),$$

where U_{fb} represents the whole feedback-loop. Just applying the definition of the MGF we obtain

$$\begin{aligned} & \Phi_{U_{sys}(s,t)}(-\theta) \\ &= \mathbb{E}(e^{-\theta \wedge (U_{fb})^{(n)} \otimes U(s,t)}) = \mathbb{E}\left(\bigvee_{n=0}^{\infty} e^{-\theta U_{fb}^{(n)} \otimes U(s,t)}\right). \end{aligned}$$

A naive approach would be to use $\mathbb{E}(X \vee Y) \leq \mathbb{E}(X) + \mathbb{E}(Y)$ for some positive random variables X and Y , resulting in:

$$\Phi_{U_{sys}(s,t)}(-\theta) \leq \sum_{n=0}^{\infty} \mathbb{E}(e^{-\theta U_{fb}^{(n)} \otimes U(s,t)})$$

This however is problematic, since finite representations as in (2.0.3) for $U_{fb}^{(n)}$ with $n \rightarrow \infty$, are hard to achieve, even if U_{fb} consists of a single server U and the window element Σ only.

Remembering condition (3.0.9) we choose another path here. To fix notations, insert for the placeholder in Figure 2.0.1 a dynamic V -server, such that $U_{fb}(s, t) = U \otimes V(s, t) + \Sigma(t)$. We saw in the univariate case for $b < \Sigma$ an easy solution to the feedback inequality. The same holds for the bivariate setting with a dynamic window; assume it holds for all $s \leq t$:

$$(6.0.1) \quad U \otimes V(s, t) - (U \otimes V)^{(2)}(s, t) \leq b \leq \Sigma_{min}(t)$$

with $\Sigma_{min}(t) := \min_{s \leq t} \{\Sigma(s)\}$. Then we have:

$$\begin{aligned} & U_{fb} \otimes U_{fb}(s, t) \\ &= \min_{s \leq r \leq t} \{U \otimes V(s, r) + \Sigma(r) + U \otimes V(r, t) + \Sigma(t)\} \\ &\geq \min_{s \leq r \leq t} \{U \otimes V(s, r) + U \otimes V(r, t)\} + \Sigma_{min}(t) + \Sigma(t) \\ &\geq U \otimes V(s, t) + (\Sigma_{min}(t) - b) + \Sigma(t) \\ &\geq U_{fb}(s, t) \end{aligned}$$

and U_{fb} is subadditive.

We use this property to achieve a service description for the whole system by:

$$\begin{aligned}
& U_{sys}(s, t) \\
&= \overline{U_{fb}} \otimes U(s, t) = (\mathbf{1} \wedge U_{fb}) \otimes U(s, t) \\
&= U(s, t) \wedge \min_{s \leq r \leq t} \{(U \otimes V) \otimes U(s, r) + \Sigma(r) + U(r, t)\} \\
(6.0.2) \quad & \geq U(s, t) \wedge (U \otimes V) \otimes (U \otimes V)(s, t) + \Sigma_{min}(t) \\
& \stackrel{(\text{??})}{\geq} U(s, t) \wedge U \otimes V(s, t) - b + \Sigma_{min}(t) \\
(6.0.3) \quad & \geq U(s, t) \wedge U \otimes V(s, t) \\
& = U \otimes V(s, t).
\end{aligned}$$

In (6.0.2) and (6.0.3), we used the monotonicity of min-plus convolution: $U(s, t) \geq U \otimes V(s, t)$ for any V with $V(t, t) = 0$ for all t . So, under the assumption of (6.0.1) we obtain the same service for the throttled system as for an unthrottled one where the servers U and V would have to be traversed. Hence, we are interested in the probability of (6.0.1) happening and call that event E . With this information at hand we can analyse the whole system by:

$$\begin{aligned}
& \mathbb{P}(d_{sys}(t) > x) \\
&= \mathbb{P}(d_{sys}(t) > x \mid E)\mathbb{P}(E) + \mathbb{P}(d_{sys}(t) > x \mid \neg E)\mathbb{P}(\neg E) \\
&\leq \mathbb{P}(d_{U \otimes V}(t) > x) + \mathbb{P}(\neg E)
\end{aligned}$$

Where $\mathbb{P}(d_{U \otimes V} > x)$ can be calculated by applying Theorem 3 and Theorem 8.

We now discuss condition (6.0.1) and its corresponding probability. For ease of notation, we leave the placeholder blank, i.e., $U \otimes V = U$.

One can rewrite (6.0.1) by:

$$\max_{0 \leq s \leq t} \{U(s, t) - U \otimes U(s, t)\} \leq b,$$

which is just the expression we arrive at when bounding the buffer in bivariate deterministic network calculus, if feeding a “flow” U in a dynamic $U \otimes U$ -server. Such stochastically dependent systems can be analysed in MGF-Calculus by using Hölder’s inequality. For this assume some MGF-bounds of the form (2.0.2) and (2.0.3) for U (we denote the σ and ρ corresponding to bound (2.0.2) by $\bar{\sigma}$ and $\bar{\rho}$ to distinguish them from the ones used in (2.0.3)). The probabilistic backlog bound

in such a system is $(\frac{1}{p} + \frac{1}{q} = \frac{1}{p'} + \frac{1}{q'} = 1)$:

$$\begin{aligned}
\mathbb{P}(q(t) > b) &\leq \mathbb{P}(\max_{0 \leq s \leq t} \{U(s, t) - U \otimes U(s, t)\} > b) \\
&\leq e^{-\theta b} \sum_{s=0}^t \mathbb{E}(e^{\theta(U(s, t) - U \otimes U(s, t))}) \\
&\leq e^{-\theta b} \sum_{s=0}^t (\mathbb{E}(e^{p\theta U(s, t)})^{1/p} (\mathbb{E}(e^{-\theta q U \otimes U(s, t)})^{1/q}) \\
&\leq e^{-\theta b} \sum_{s=0}^t e^{\theta \bar{\rho}(p\theta)(t-s) + \theta \bar{\sigma}(p\theta)} (\mathbb{E}(e^{-\theta q U \otimes U(s, t)})^{1/q}) \\
&\leq e^{-\theta b + \theta \bar{\sigma}(p\theta)} \sum_{s=0}^t e^{\theta \bar{\rho}(p\theta)(t-s)} \\
&\quad \cdot \left(\sum_{r=s}^t e^{\theta q \rho(qp'\theta)(r-s) + \theta q \sigma(qp'\theta)} e^{\theta q \rho(qq'\theta)(t-r) + \theta q \sigma(qq'\theta)} \right)^{1/q} \\
&\leq e^{-\theta b + \theta \bar{\sigma}(p\theta) + \theta \sigma(qp'\theta) + \theta \sigma(qq'\theta)} \\
&\quad \cdot \sum_{s=0}^t e^{\theta \bar{\rho}(p\theta)(t-s)} e^{\theta \rho(qp'\theta)(t-s)} \left(\sum_{r'=0}^{t-s} e^{\theta q(\rho(qq'\theta) - \rho(qp'\theta))r'} \right)^{1/q}
\end{aligned}$$

We have used Chernoff's inequality in the second line and Hölder's inequality in line 3 and 5. If we assume, w.l.o.g. that $q' > p'$ holds, we have that the last sum is convergent in t . We denote the limit of that series just by B and can proceed with:

$$\begin{aligned}
\mathbb{P}(q(t) > b) &\leq e^{-\theta b + \theta(\bar{\sigma}(p\theta) + \sigma(qp'\theta) + \sigma(qq'\theta))} B \sum_{s=0}^t e^{\theta \bar{\rho}(p\theta)(t-s) + \theta \rho(qp'\theta)(t-s)} \\
&\leq e^{-\theta b + \theta(\bar{\sigma}(p\theta) + \sigma(qp'\theta) + \sigma(qq'\theta))} B \sum_{s'=0}^t e^{\theta(\bar{\rho}(p\theta) + \theta \rho(qp'\theta))s'}.
\end{aligned}$$

The above sum, however, does in general not converge. This is due to the three inequalities:

$$\begin{aligned}
\bar{\rho}(\theta) &\geq -\rho(\theta) \quad \forall \theta \geq 0, \\
\bar{\rho}(p\theta) &\geq \bar{\rho}(\theta) \quad \forall p \geq 1, \\
\rho(q\theta) &\leq \rho(\theta) \quad \forall q \geq 1.
\end{aligned}$$

For the case $q' = p' = 2$, the exponents in the last sum vanish and the expression reduces to $t - s + 1 \leq e^{\theta(t-s)\frac{1}{\theta}}$, such that the bound becomes:

$$\begin{aligned}
\mathbb{P}(q(t) > b) &\leq e^{-\theta b + \theta(\bar{\sigma}(p\theta) + 2\sigma(2q\theta))} \sum_{s'=0}^t e^{\theta(\bar{\rho}(p\theta) + \theta \rho(2q\theta) + \frac{1}{q\theta})s'}.
\end{aligned}$$

Note that the system $\xrightarrow{U} U \rightarrow U$ is unstable (the concatenation of the two servers with service $U \otimes U$ is at most as large as the arrivals U), which is why the above bound is valid only for any finite time t .

Overall, we can summarize our findings in the main result of this work:

THEOREM 26. *Consider a WFC system as in Figure 2.0.1 with the placeholder being a dynamic V -server. Let $\frac{1}{p} + \frac{1}{q} = \frac{1}{p'} + \frac{1}{q'} = 1$ and $q' > p'$. Assume the following MGF-bounds on $U \otimes V$:*

$$\begin{aligned}\mathbb{E}(e^{-\theta U \otimes V(s,t)}) &\leq e^{\theta \rho(\theta)(t-s) + \sigma(\theta)} \\ \mathbb{E}(e^{\theta U \otimes V(s,t)}) &\leq e^{\theta \bar{\rho}(\theta)(t-s) + \bar{\sigma}(\theta)}.\end{aligned}$$

The whole system fulfills the probabilistic delay-bound:

$$\begin{aligned}\mathbb{P}(d_{sys}(t) > x) &\leq \mathbb{P}(d_{U \otimes V}(t) > x) \\ &+ e^{-\theta \Sigma_{min}(t+T) + \sigma_E(\theta, p, p')} B \sum_{s=0}^{t+T} e^{\theta \rho_E(\theta, p, p')s}\end{aligned}$$

with

$$\begin{aligned}\sigma_E(\theta, p, p') &= \bar{\sigma}(p\theta) + \sigma(qp'\theta) + \sigma(qq'\theta) \\ \rho_E(\theta, p, p') &= \begin{cases} \bar{\rho}(p\theta) + \rho(qp'\theta) & \text{if } p' \neq 2 \\ \bar{\rho}(p\theta) + \rho(2q\theta) + \frac{1}{q\theta} & \text{if } p' = 2 \end{cases} \\ B &= \begin{cases} \left(\frac{1}{1 - e^{\theta q(\rho(qq'\theta) - \rho(qp'\theta))}} \right)^{1/q} & \text{if } p' \neq 2 \\ 1 & \text{if } p' = 2 \end{cases}\end{aligned}$$

and $d_{U \otimes V}$ being the delay of an unthrottled tandem consisting of a dynamic U - and V -server.

We want to emphasize that with this theorem it is for first time possible to analyse a general WFC system in the context of SNC. Our solution does not rely on the subadditivity of U or V directly, in contrast to what was presented in Chapter 3 and 5. Instead we ask for the probability of failing the subadditivity by at least the window size Σ , which allows an analysis of general service elements inside the feedback loop. Note that U and V do not need to be single service elements themselves, they could instead result from Theorem 3 or include further elements (like fixed delay-elements or scaling elements [9]).

REMARK 27. It is interesting to note, that, if $U \otimes V$ is subadditive already, we have $\mathbb{P}(\neg E) = 0$ and the whole system's service reduces immediately to

$$U_{sys}(s, t) = U(s, t) \wedge U \otimes V(s, t) + \Sigma_{min}(t).$$

One can directly apply Theorem 8 on U_{sys} to achieve an end-to-end delay bound in this case. Further, we observe that step (6.0.3) is not a necessary one. We could, for example, shift b below $\Sigma_{min}(t)$ and continue directly with:

$$(6.0.4) \quad U_{sys}(s, t) \geq U(s, t) \wedge U \otimes V(s, t) + \Sigma_{min}(t) - b.$$

We can then apply Theorem 26 on this U_{sys} (which is at least as large as $U \otimes V$). One can view this as a shift in the violation probabilities towards the subadditive part (event E) of the bound. We investigate this tradeoff in the following section.

6.1. Numerical Evaluation

In this section, we investigate how the bound derived in Theorem 26 evolves in its parameters. Further, we quantify the impact of WFC on the delay of the system, by comparing it to a similar unthrottled system. For this section we assume U to be a constant rate server $U_o(s, t) = u(t - s)$, which also serves a crossflow $A_U(s, t)$ at higher priority than B . A well-known result in SNC states that B receives a service $U(s, t) = u(t - s) - A_U(s, t)$. Similarly, we insert for the placeholder in Figure 2.0.1 a service element V , which also offers a constant rate $V_o(s, t) = v(t - s)$ shared with a higher priority crossflow $A_V(s, t)$, such that $V(s, t) = v(t - s) - A_V(s, t)$. Note that both service descriptions are subadditive by themselves, but when applying Theorem 3 this property is lost.

To account for the typically smaller size of acknowledgements flowing back to the throttle, we assume $v > u$. The crossflows in this example consist of i.i.d. exponentially distributed increments $a_U(t)$ and $a_V(t)$, respectively. The arrivals to the WFC system, denoted by A also consist of i.i.d. exponentially distributed increments. All flows are stochastically independent of each other. More sophisticated crossflows or arrivals are possible to analyse, as well as dropping the independency assumption, yet this is not the focus of our evaluation. We further assume a constant window-size Σ , for all times $t \geq 0$.

A corresponding unthrottled system would just consist of the flow A being fed into the service element U , thus Theorem 8 could be applied directly.

To achieve reasonable values in the bounds of Theorem 8 and Theorem 26, we numerically optimized the parameter θ and the Hölder-pairs p, q, p', q' . If not specified otherwise we used the following set of parameters in our calculations: the bound is taken at time $t = 5$ and asks for a delay $T = 10$, i.e., we consider the probability $\mathbb{P}(d_{sys}(5) > 10)$. The parameter of the exponential distributions for the arrivals and cross-flows is given by $\lambda = 4$ (we assume all three flows to have the same rate λ for simplicity), while the server-rates are $u = 1$ and $v = 2$. This corresponds to a utilization of 50% and 25%, respectively. We present the results for a window size of $\Sigma = 15$.

6.1.1. Throttled vs. Unthrottled System. First we want to compare the system to its unthrottled counterpart. To that end, we alter the arrival rates λ , resulting in utilizations from 30% to 80%. We plot the corresponding violation probabilities for the performance bounds on a logarithmic scale for the throttled, as well as the unthrottled system. We did this for different window-sizes $\Sigma = 10, 15, 20$. The results are displayed in Figure 6.1.1 as black and blue lines for the throttled and unthrottled system, respectively. As expected, the throttled system behaves better, the larger Σ is; for $\Sigma = 20$ the throttled system behaves almost identically to the unthrottled one.

6.1.2. Dependence on Delay. A major difference between the unthrottled and the throttled analysis lies in the dependency on the delay T . While for the unthrottled system an increase in T leads to a decrease in the violation probability, we see in the bound of Theorem 26, that the term $\mathbb{P}(\neg E)$ increases in T . In Figure (6.1.2), one sees for the black line how the bound evolves for an increasing T . The two red lines show how the bound differs when choosing $b = \frac{\Sigma}{2}, \frac{9\Sigma}{10}$ as suggested in Remark 27 (Equation (6.0.4)). The trend here for larger T is, that the bound becomes worse, the larger the difference between b and Σ is. However, for small

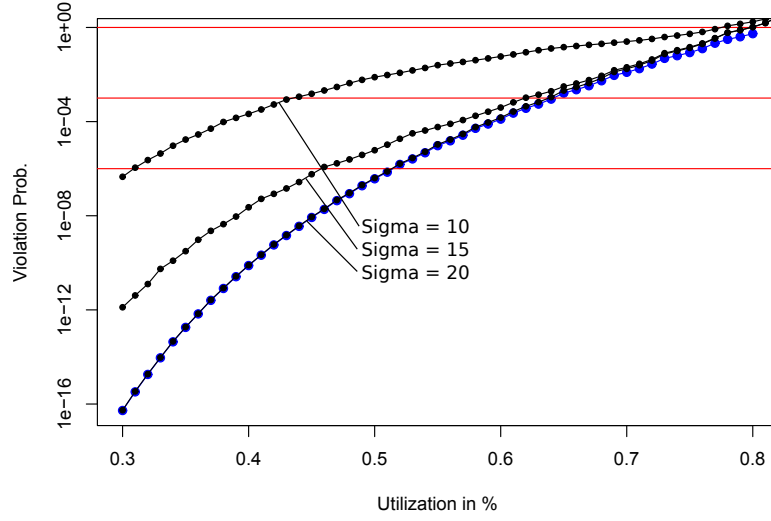


FIGURE 6.1.1. A graph showing the violation probabilities depending on the utilization of throttled (black) and unthrottled (blue) systems for different window sizes. The red lines are equal to 1, 10^{-3} and 10^{-6} .

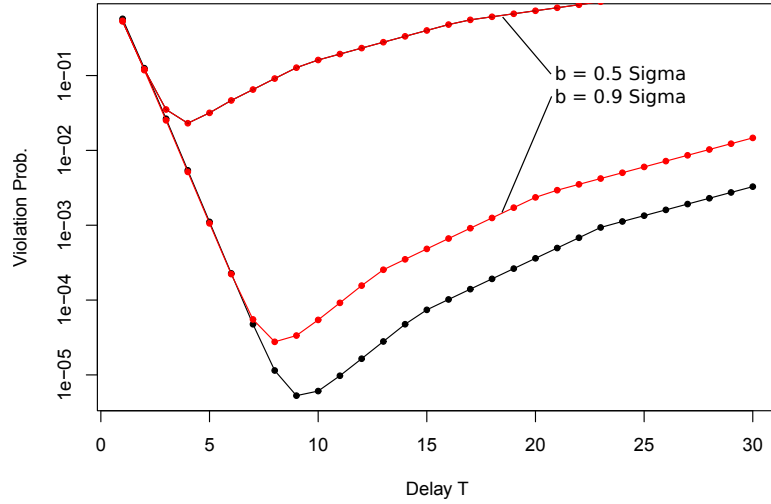


FIGURE 6.1.2. A graph showing how the bound evolves when increasing the delay T . The red colored lines represent a shift towards the violation probability of event E .

values of T there is a very slight improvement for $b = \frac{9\Sigma}{10}$ and even for $b = \frac{\Sigma}{2}$. In this scenario, trading a higher violation probability for the event E is not worthwhile the gain from a better service description U_{sys} .

To investigate the composition of the delay bound further, we separated the two parts of the bound in Figure (6.1.3). The blue circles correspond to the delay-part of the violation probability $\mathbb{P}(d_{U \otimes V}(t) > x)$ and the red circles correspond to

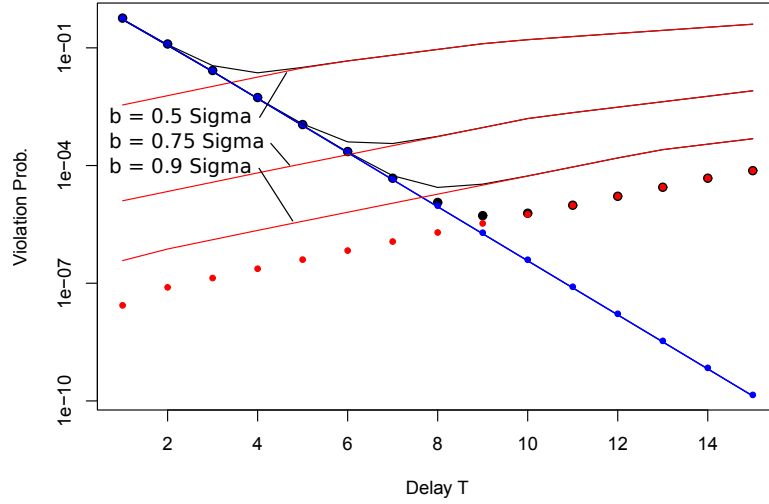


FIGURE 6.1.3. A graph showing the different components of the violation probability: the blue circles are the delay-part, while the red circles represent the subadditivity-part. The black circles are the sum of both parts. The lines show the same for different shifts towards the violation probability $\mathbb{P}(\neg E)$.

the violation probability of event E , while the solid black circles are the sum of both. It can be clearly observed that from a certain point onwards the probability $\mathbb{P}(\neg E)$ dominates the overall violation probability. The additional lines drawn into the graph, show how the different parts of the bound are affected when we use $b = \frac{\Sigma}{2}, \frac{3\Sigma}{4}, \frac{9\Sigma}{10}$ in Equation (6.0.4). It can be seen that the delay-part (blue) of the probability experiences no considerable change, while the probability of violating event E (red) increases significantly, when $b < \Sigma$.

6.1.3. Convergence to Unthrottled System. In Figure (6.1.4), we consider the convergence of the throttled system towards the unthrottled one when increasing the window size. Clearly, from Theorem (26) the violation probability $\mathbb{P}(\neg E)$ vanishes for increasing window sizes. However, the throttled system does not fully converge to the unthrottled one, since the delay-part of the bound still differs:

$$\mathbb{P}(d_U(t) > T) \leq \mathbb{P}(d_{U \otimes V}(t) > T).$$

The size of the gap, which cannot be closed by increasing the window size further is completely dependent on the service descriptions U and V . We present in the graph the same system as before, but vary $v = 2, 1.5, 1.1$. One can see clearly how the gap to the delay of the unthrottled system (red) increases, when reducing the rate of v .

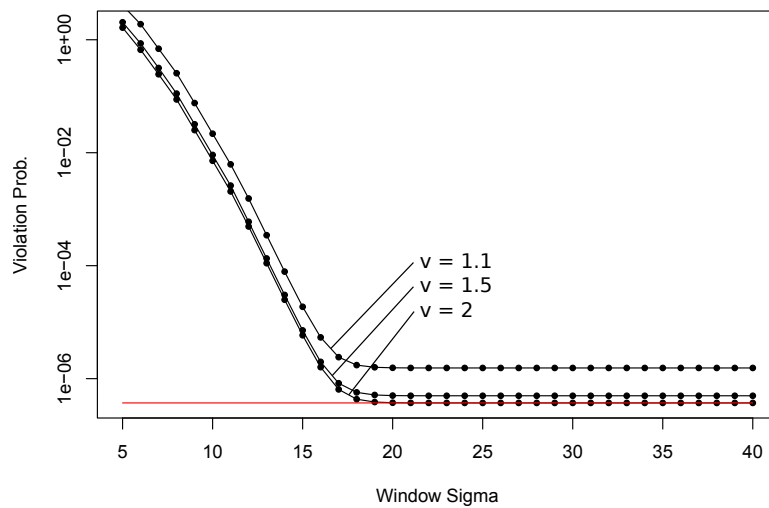


FIGURE 6.1.4. A graph showing the convergence of the throttled system towards the unthrottled one, when increasing the window Σ . The red line is the bound for the unthrottled system. The black lines show the throttled system, for different rates of v .

Conclusion and Outlook

In this work, we have dealt with the long-standing problem of analysing WFC systems in SNC. While such feedback loops had been solved in deterministic network calculus more than a decade ago, its counterpart in the stochastic setting has been a well-known open problem [14, 11, 12, 7]. We presented how far subadditive service carries DNC solutions for WFC systems into stochastic network calculus. In that discussion, we encountered the very general notion of σ -additive operators and saw as a tractable example a feedback-loop containing a single subadditive server. Unfortunately, this method reaches the end of the road as soon as operators appear which no longer commute, or are not idempotent. This is not untypical in applications, for example if tandems of servers are involved.

Therefore, we approached the problem in two different ways and hence, for the first time, successfully analysed general WFC systems in the context of SNC. For the first method, we identified subadditivity and its weaker defective version as key for tackling WFC in SNC. Building upon this property we gave the first stochastic performance bounds in WFC systems. Although the assumption of subadditive service is rather strong, we were able to identify feed-forward networks with strict priority scheduling as subadditive networks. Future work on this method includes focusing on more general topologies or schedulers (e.g. FIFO).

In the second method instead of assuming subadditive service elements, we leverage the stochastic nature of the problem and ask for the probability of the feedback loop not being subadditive. This effectively allows the analysis of WFC systems in MGF-based network calculus. The resulting bounds consist of two parts: first, a delay-bound of a conventional unthrottled system, containing the feedback loop as service; second, a probability of violating the subadditivity, by more than the window size. The structure of our result makes a direct comparison between throttled and unthrottled systems possible. The presented method uses a backlog bound for the system $\xrightarrow{U} U \otimes U \rightarrow$. The “arrivals” and “service” in this scenario are strongly correlated. While using Hölder’s inequality deals correctly with that dependence, it also neglects its possible advantages. As the arrivals and the service in this system are positively correlated one can hope to improve the bounds significantly, when taking the dependencies into account.

The analysis of WFC systems in stochastic network calculus is not completed yet, but has rather just begun. While the now available methods can handle varying window sizes Σ , they can take only limited advantage of their variations.

Besides improving both methods with respect to tightening the bounds, one can extend and built upon this work: one direction is to break the “end-to-end” feedback-loop into several hops, resulting in a tandem of WFC systems. Another interesting question would be how to effectively handle stochastic dependencies

between the “upstream”-service U , the “down-stream”-elements and the window-process Σ . Answering this will push the applicability of SNC even further. By better grasping the occurring dependencies one can eventually aim at analysing systems like the window-controlled TCP in SNC.

Bibliography

- [1] R. Agrawal, R. L. Cruz, C. Okino, and R. Rajan. Performance bounds for flow control protocols. *IEEE/ACM Transactions on Networking*, 7(3):310–323, June 1999.
- [2] M. A. Beck. A first course in stochastic network calculus. Online lecture notes, October 2013. <http://goo.gl/qSIIwr>.
- [3] J.-Y. Le Boudec and P. Thiran. *Network Calculus - A Theory of Deterministic Queuing Systems for the Internet*. Springer, 2001.
- [4] C.-S. Chang. On deterministic traffic regulation and service guarantees: A systematic approach by filtering. *IEEE Transactions on Information Theory*, 44(3):1097–1110, May 1998.
- [5] C.-S. Chang. *Performance Guarantees in Communication Networks*. Springer, 2000.
- [6] F. Ciucu, A. Burchard, and J. Liebeherr. A network service curve approach for the stochastic analysis of networks. In *Proc. of ACM SIGMETRICS*, June 2005.
- [7] F. Ciucu, M. Fidler, J. Liebeherr, and J. Schmitt. Network calculus. *Dagstuhl Reports*, pages 63–83, March 2015.
- [8] F. Ciucu and J. Schmitt. Perspectives on network calculus – no free lunch, but still good value. In *Proc. of ACM SIGCOMM*, August 2012.
- [9] F. Ciucu, J. Schmitt, and H. Wang. On expressing networks with flow transformations in convolution-form. In *Proc. of IEEE INFOCOM*, April 2011.
- [10] M. Fidler. An end-to-end probabilistic network calculus with moment generating functions. In *Proc. of IWQoS*, June 2006.
- [11] M. Fidler. Survey of deterministic and stochastic service curve models in the network calculus. *IEEE Communications Surveys Tutorials*, 12(1):59–86, 2010.
- [12] M. Fidler and A. Rizk. A guide to the stochastic network calculus. *IEEE Communications Surveys and Tutorials*, 17(1):92–105, 2015.
- [13] Y. Ghiassi-Farrokhfal, S. Keshav, and C. Rosenberg. Toward a realistic performance analysis of storage systems in smart grids. *IEEE Transactions on Smart Grid*, 6(1):402–410, Jan 2015.
- [14] Y. Jiang and Y. Liu. *Stochastic Network Calculus*. Springer, 2008.
- [15] J.-Y. Le Boudec and P. Thiran. A note on time and space methods in network calculus. In *International Zurich Seminar on Broadband Communications*, pages 267–272, Feb 1998.
- [16] J. Schmitt, F.A. Zdarsky, and L. Thiele. A comprehensive worst-case calculus for wireless sensor networks with in-network processing. In *Proc. of IEEE RTSS*, pages 193–202, Dec 2007.