Network-Layer Performance Analysis of Multi-Hop Fading Channels

Hussein Al-Zubaidy, Jörg Liebeherr, Almut Burchard

***Abstract*—A fundamental problem for the delay and backlog analysis across multi-hop paths in wireless networks is how to account for the random properties of the wireless channel. Since the usual statistical models for radio signals in a propagation environment do not lend themselves easily to a description of the available service rate, the performance analysis of wireless networks has resorted to higher-layer abstractions, e.g., using Markov chain models. In this work, we propose a network calculus that can incorporate common statistical models of fading channels and obtain statistical bounds on delay and backlog across multiple nodes. We conduct the analysis in a transfer domain, where the service process at a link is characterized by the instantaneous signal-to-noise ratio at the receiver. We discover that, in the transfer domain, the network model is governed by a dioid algebra, which we refer to as the $(\min,\times)$ algebra. Using this algebra we derive the desired delay and backlog bounds. Using arguments from large deviations theory, we show that the bounds are asymptotically tight. An application of the analysis is demonstrated for a multi-hop network of Rayleigh fading channels with cross traffic at each hop.**

I. Introduction

Network-layer performance analysis seeks to provide estimates on the delays experienced by traffic traversing the elements of a network, as well as the corresponding buffer requirements. For wireless networks, a question of interest is how the stochastic properties of wireless channels impact delay and backlog performance. Wireless channels are characterized by rapid variations of the channel quality caused by the mobility and location of communicating devices. This is due to *fading*, which is the deviation in the attenuation experienced by the transmitted signal when traversing a wireless channel. The term *fading channel* is used to refer to a channel that experiences such effects. In this paper we explore the network-layer performance of a multi-hop network where each link is represented by a fading channel.

We model the multi-hop wireless network by tandem queues with randomly varying capacity servers, where each server represents the random capacity of a fading channel. We assume that the transmission rates over the fading channels are equal to their information-theoretic capacity limit in bits per second, C. The capacity limit is expressed as a function of the instantaneous signal-to-noise ratio (SNR) at the receiver, γ, by $C(\gamma) = W \log_2(1+\gamma)$, where W is the channel bandwidth (in *Hz*). This model assumes the existence of a channel coding scheme that achieves the channel capacity with arbitrarily small error probability. It does not take into consideration the relationship between codeword length and available link capacity, i.e., that transmission of long codewords requires additional link capacity, whereas short codewords induce transmission errors that trigger retransmissions. Numerous models are available to describe the gain of fading channels depending on the type of fading (slow or fast), and the environment (e.g., urban or rural). We express the capacity of a fading channel as the natural logarithm of a function of γ by (see Chapter 14.2 in [30])

$$C(\gamma) = c \log g(\gamma)\,, \tag{1}$$

where c is a constant and $g(\gamma)$ is used to characterize the fading channel. We are interested in finding bounds on the end-to-end delay and backlog for a cascade of fading channels, with store-and-forward processing at each channel.

The analysis in this paper follows a system-theoretic stochastic network calculus approach [18], which describes the network properties using a $(\min,+)$ dioid algebra. Arrival and departure processes at a network element are described by bivariate stochastic processes $A(\tau,t)$ and $D(\tau,t)$, respectively, denoting the cumulative arrivals and departures in the time interval $[\tau,t)$. A network element is characterized by the service process $S(\tau,t)$, denoting the available service in $[\tau,t)$. The input-output relationship at the network element is described by

$$D(0,t) \geq A * S(0,t)\,, \tag{2}$$

where the $(\min,+)$ convolution operator '$*$' is defined as $f * g(\tau,t) = \inf_{\tau \leq u \leq t}\{f(\tau,u) + g(u,t)\}$. If network traffic passes through a tandem of N network elements with service processes $S_1, S_2, \ldots, S_N$, then the traffic will see a network service process that is given by $S_1 * S_2 * \ldots * S_N$.

The stochastic properties of fading channels present a challenge for a network-layer analysis since the service processes corresponding to the channel capacity of common fading channel models, such as Rician, Rayleigh, or Nakagami-m, require to take a logarithm of their distributions. As discussed in the next section, researchers frequently turn to higher-layer abstractions in order to overcome the difficulty of working directly with these distributions. Widely used abstractions are the two-state Gilbert-Elliott model and its extension to a finite-state Markov channel (FSMC) [31]. FSMC models simplify the analysis to a degree that the network model becomes tractable, at least at a single node. Extensions to multi-hop settings encounter a rapidly growing state space. As of today, a

H. Al-Zubaidy and J. Liebeherr are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 1A1 Canada, e-mail: {hzubaidy, jorg}@comm.utoronto.ca.

A. Burchard is with the Department of Mathematics, University of Toronto, e-mail: almut@math.toronto.edu. This manuscript is an extended and revised version of an Infocom conference publication [1].

general multihop analysis that is applicable to common fading channel models remains an open problem.

In this paper, we pursue a novel approach to the analysis of multi-hop wireless networks. We develop a calculus for wireless networks that utilizes a service description in terms of a physical-layer fading channel model and a traffic description based on network-layer traffic models in order to provide end-to-end performance bounds. We view the network-layer model with arrival, departure and service processes as residing in a *bit domain*, where traffic and service is measured in bits. We view the fading channel models used in wireless communications as residing in an alternate domain, which we call the *SNR domain*, where channel properties are expressed in terms of the distribution of the signal-to-noise ratio at the receiver. We then derive a method to compute performance bounds from these traffic and service characterizations.

A key observation in our work is that service elements in the SNR domain obey the laws of a dioid algebra. We devise a suitable dioid, the $(\min, \times)$ *algebra*, where the minimum takes the role of the standard addition, and the second operation is the standard multiplication, and use it for analysis in the SNR domain. In this domain, multi-hop descriptions of fading channels become tractable. In particular, we find that a cascade of fading channels can be expressed in terms of a convolution in the new algebra of the constituting channels. The key to our analysis is that we derive performance bounds entirely in the SNR domain. Observing that the bit and SNR domains are linked by the exponential function, we transfer arrival and departure processes from the bit to the SNR domain. Then, we derive backlog and delay bounds in the transfer domain using the $(\min, \times)$ algebra. The results are mapped back to the original bit domain to finally give us the desired performance bounds. Using results from large deviations theory, we show that our bounds are asymptotically tight. Our derivations in the SNR domain require the computation of products and quotients of random variables. Here, we use the *Mellin transform* to facilitate otherwise cumbersome calculations. Then, the computational problem is reduced to finding the Mellin transform for service and traffic processes. Although, the mapping from $(\min, +)$ algebra to $(\min, \times)$ algebra is one-to-one, the $(\min, \times)$ algebra is better suited for the analysis of wireless networks performance.

The main contribution of this paper is the development of an alternative approach for modeling the impact of channel gain models on the network-layer performance of wireless networks. For the purposes of this paper, the SNR domain is used solely as a transfer domain that enables us to solve an otherwise intractable mathematical problem. On the other hand, the ability to relate quantities that appear in network-layer models and concepts found in a physical-layer analysis may prove useful in a broader context, e.g., for studying cross-layer performance issues in wireless communications. Moreover, the $(\min, \times)$ algebra and the Mellin transform form a tool set that can be applied more generally in wireless communications for studying the channel gain of cascades of fading channels.

We emphasize that this paper only considers simple network scenarios and makes numerous convenient assumptions (which are made explicit in Sec. III).

In addition to the channel capacity model, which ignores the effects of finite code block lengths, we assume a time slotted system where the signal-to-noise ratio in each time slot is independent and identically distributed. This assumption is justified when the sampling intervals for the channel state are longer than the channel coherence time. We also assume that fading remains constant during the length of the sampling interval. A channel model that conforms to these assumptions is block fading without memory [29]. Frequency hopping (FH) is a system where such a channel model applies [19]. This results in an idealized analytic model of the wireless channel. Refinements of the model that account for finite codeword lengths, imperfect coding and decoding, and fading channels with memory, all of which result in a discounted service process, are left for future work. We emphasize that the network calculus methodology can be applied to settings without independence, i.e., channels with memory. This requires additional model parameters to characterize the time correlations. Relaxations of other assumptions, as well as extensions of the presented model using a network calculus methodology remain open research problems.

The remainder of the paper is organized as follows. In Sec. II we discuss related work. We describe the system model in Sec. III, where we also motivate the use of the SNR domain. In Sec. IV we present the $(\min, \times)$ algebra and derive performance bounds. We also address the tightness of these bounds. In Sec. V we apply the analysis to a cascade of Rayleigh channels. In Sec. VI we present numerical examples. Conclusions are provided in Sec. VII.

II. Related Work

Approaches for network-layer performance analysis of wireless networks include queueing theory, effective bandwidth and, more recently, network calculus. Since the service processes corresponding to the channel capacity of Rician, Rayleigh, or Nakagami-m fading channel models require to take a logarithm of their distributions, researchers often turn to higher-layer abstractions to model fading channels, which lend themselves more easily to an analysis. Widely used abstractions are the two-state Gilbert-Elliot channel model and subsequent extensions to a finite-state Markov channel (FSMC) [36]. Markov channel models are well suited to express the time correlation of fading channel samples.

Queueing theoretic studies of fading channels generally apply approximations to reduce the complexity of multi-hop models. Le, Nguyen, and Hossain [22] pursue a decomposition approximation to analyze the loss probability and average delay of a multi-hop wireless network with slotted transmissions for a batch Bernoulli arrival process, and with independent cross traffic at each node. Another decomposition approximation is presented by Le and Hossain [21], who consider a multi-hop tandem network with a batch arrival process and multi-rate transmissions, to develop a routing scheme that can meet given delay and loss requirements. The analysis obtains end-to-end loss rates and delays with a decomposition analysis, and feeds the results as metrics to the routing algorithm. Bisnik and Abouzeid [4] model a multi-hop wireless network as an open network of G/G/1 queueing systems. Using diffusion approximation, they obtain closed-form expressions for average end-to-end delays.

Effective bandwidth [20] analysis seeks to develop (asymptotic) bounds on performance metrics, e.g., an exponential decay of the backlog. Wu and Negi [39] have adapted an effective bandwidth approach to the analysis of fading channels. They introduce the concept of effective capacity, which characterizes a wireless channel by a log-moment generating function (log-MGF) of the channel capacity. They obtain an asymptotic approximation of the delay bound violation probability of a Rayleigh fading channel. Due to the difficulty of computing the moment generating function (MGF) of the Rayleigh distribution, they simplify the analysis by assuming non-correlated distributions with low SNR and estimate channel parameters from measurements. The work has been extended to correlated Rayleigh and correlated Nakagami-m channels, and to cascades of fading channels [37], [38], [40]. A closely related concept is the effective channel capacity presented by Li et al. [25], which describes the available channel capacity by a first order Markov chain and computes the log-MGF of the underlying Markov process. Using methods developed in [24], they compute statistical delay bounds for Nakagami-m fading channel. Hassan, Krunz, and Matta [16] use an effective bandwidth analysis to study delay and loss performance at a single wireless link, which is modeled by an FSMC. For fluid On-Off traffic and FIFO buffering, they obtain a closed form expression for the effective bandwidth required to guarantee bounds on delay and packet loss.

There is a collection of recent works that apply stochastic network calculus methods to wireless networks with fading channels. The stochastic network calculus is closely related to the effective bandwidth theory, in that it seeks to develop bounds on performance metrics under assumptions also found in the effective bandwidth literature [18]. Different from the effective bandwidth literature, stochastic network calculus methods seek to develop non-asymptotic bounds. An attractive element of a network calculus analysis is that it is often possible to extend a single node analysis to a tandem of nodes, using the $(\min, +)$ convolution.

Fidler [13] presents a network calculus methodology for a two-state FSMC model of a single-hop fading channel. He applies the MGF network calculus from [6, Chapter 7], [12]. The MGF network calculus takes its name from the extensive use of moment generating functions in the derivation of performance bounds. Mahmood, Rizk, and Jiang [27] apply the MGF network calculus to MIMO channels and derive delay bounds for periodic traffic sources. Zheng et. al. [41] also use an MGF network calculus to study the performance of two-hop relay networks. A similar methodology is applied in [28] to compute the throughput of a multi-user DS-CDMA system with delay constraints and in [42] to study the performance of a wireless finite-state Markov channel. In the MGF network calculus based work above, models for a cascade of fading channels become complex, so that multi-node results for networks with more than two nodes have not been obtained.

The $(\min, \times)$ network calculus developed in this paper uses similar descriptions and assumptions for traffic and service as the MGF network calculus. By performing computations in a transfer domain, where fading channel models take a simpler form, we are able to compute multi-node service descriptions for an arbitrarily large number of nodes.

The MGF network calculus assumes that arrivals and service at each node are independent. These assumptions can be relaxed using statistical envelope descriptions for traffic (effective envelopes) and service (statistical service curve) [5], [18]. Jiang and Emstad [17] have applied an approach with envelopes to a fading channel that is characterized by two stochastic processes: an ideal service process and an impairment process, where the impairment process captures effects due to fading, noise, and cross traffic. Verticale and Giacomazzi [35] have obtained a closed form expression for the variance of a service curve that describes the available service by a Markov chain. This is used for the analysis of an FSMC model of a Rayleigh fading channel. For computing the bounds for Markovian arrivals, they apply the bounded-variance network calculus introduced in [14], which is an extension of the central limit theorem methods by Choe and Shroff [8] to multi-hop paths. Verticale [34] has applied the same methodology to constant bit rate traffic. Ciucu, Pan, and Hohlfeld [10] and Ciucu [9] present closed-form expressions for the delay and throughput distributions for multi-hop wireless networks. Here, the fading channel is modeled by an abstraction that uses a link layer model of the transmission channel. The channel is assumed to behave like a slotted-ALOHA system in half-duplex mode. The model of this channel is a two-state On-Off server, where a node can transmit (i.e., is in the On state) only when all other nodes in the interference range are not transmitting.

There is also a literature on physical-layer performance metrics of fading channels in multi-hop wireless networks. Hasna and Alouini [15] have presented a framework for evaluating the end-to-end outage probability of a multi-hop wireless relay network with independent, non-regenerative relays, i.e., amplify-and-forward (AF), over Nakagami fading channels. Similar bounds were found in [33] and [2] for the average error probability and end-to-end SNR for AF relay networks. These works study physical-layer performance bounds of channel-assisted, amplify-and-forward relaying over a multi-hop fading channels. They do not consider buffering or traffic burstiness, and are not concerned with network-layer performance metrics addressed in this paper. Delay and backlog analysis and optimization of multihop wireless networks remain open research problems [21].

III. Network Model in the Bit and SNR Domains

We consider a wireless N-node tandem network as shown in Fig. 1, where each node is modeled by a server with an infinite buffer. We are interested in the performance experienced by a (through) flow that traverses the entire network and may encounter cross traffic at each node. One can think of the cross traffic at a node as the aggregate of all traffic traversing the node that does not belong to the through flow. The service given to the through flow at a node is a random process, which

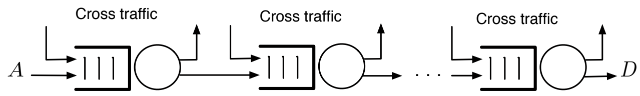

Fig. 1. Tandem network model.

is determined by the instantaneous channel capacity as well as the cross traffic at the node. We consider a fluid-flow traffic model where the flow is infinitely divisible. We will work in a discrete-time domain $\mathcal{T} = \{t_i : t_i = i \cdot \Delta t, i \in \mathbb{Z}\}$, where $\mathbb{Z}$ is the set of integers and Δt is the length of a time unit. Setting $\Delta t = 1$ allows us to replace t_i by i, which we interpret as the index of a time slot. We assume that the system is started with empty queues at time $t = 0$.

Different nodes and different traffic flows will be distinguished by subscripts. The cumulative arrivals to, service offered by, and departures from a node are represented, respectively, by random processes A, S and D that will be described more precisely below. Throughout this work, we assume that arrival and service processes satisfy stationary bounds.

A. Traffic and Service in the Bit Domain

Consider for the moment a single node. We write

$$A(\tau,t) = \sum_{i=\tau}^{t-1} a_i\,, \quad \text{and} \quad D(\tau,t) = \sum_{i=\tau}^{t-1} d_i\,,$$

for the cumulative arrivals and departures, respectively, at the node in the time interval $[\tau, t)$, where a_i denotes the arrivals and d_i the departures in the i-th time slot. Due to causality, we have $D(0,t) \leq A(0,t)$. The processes lie in the set $\mathcal{F}$ of non-negative bivariate functions $f(\tau,t)$ that are increasing in the second argument and vanish unless $0 \leq \tau < t$. The backlog at time $t > 0$ is given by

$$B(t) = A(0,t) - D(0,t)\ , \tag{3}$$

and the delay at the node is given by

$$W(t) = \inf\{u \geq 0 : A(0,t) \leq D(0,t+u)\}\ . \tag{4}$$

A node where backlog and delay increase with time and become unbounded is said to be *unstable*. Conditions to ensure that B and W are finite at all times are referred to as *stability conditions*.

The service of the node in the time interval $[\tau, t)$ is given by a random process $S(\tau,t)$, such that Eq. (2) holds for every arrival process A and the corresponding departure process D. This service description with bivariate functions is referred to as *dynamic server*. Initially defined for non-random service [7], dynamic servers have been extended to random processes in [6], [12].

The model in Fig. 1 is a classical network-layer model, where traffic is measured in bits and service is measured in bits per second. We thus refer to this model of arrivals, departures and service as residing in a *bit domain*.

The network calculus exploits that networks which satisfy the input-output relation of Eq. (2) with equality can be viewed as linear systems in a $(\min,+)$ dioid algebra [3], [23]. In the $(\mathbb{R} \cup \{+\infty\}, \min, +)$ dioid, the minimum and addition replace the standard addition and multiplication operations. The network calculus is based on the fact that $(\mathcal{F}, \min, *)$ is again a dioid [6]. Note that the min-plus convolution, which is the second operation in the dioid, is not commutative in $\mathcal{F}$.

B. Service Model for Fading Channel

We assume that the state of a wireless channel is sampled at equal time intervals. Denoting by γ_i the instantaneous signal-to-noise ratio observed at the receiver in the i-th sampling epoch, γ_i is a nonnegative random variable that has the probability distribution of the underlying fading model. We assume that the random variables γ_i are independent and identically distributed. This assumption is justified when the sampling epoch is longer than the channel coherence time. We also assume that fading remains constant for the duration of one time unit. Both assumptions are consistent with a block fading channel model. For channels that cannot be described by block fading, when choosing Δt to be longer than the coherence time, the second assumption results in a quantization error. To minimize the quantization error, Δt should be selected as the smallest value that still justifies the independence assumption.

Using Eq. (1), the instantaneous service offered by the channel in the i-th slot is given by $\log g(\gamma_i)$ and the corresponding service process is given by

$$S(\tau,t) = \sum_{i=\tau}^{t-1} \log g(\gamma_i)\,, \tag{5}$$

where, for notational simplicity, we have chosen units such that the constant in Eq. (1) takes the value $c = 1$. Computed bounds obtained with the normalization are scaled when $c \neq 1$.

The service description in Eq. (5) requires us to work with the logarithm of fading distributions, which presents a non-trivial technical difficulty via the usual network calculus or queueing theory. On the other hand, observe that the exponential $\mathcal{S}(\tau,t) = e^{S(\tau,t)}$ is described more simply by

$$\mathcal{S}(\tau,t) = \prod_{i=\tau}^{t-1} g(\gamma_i)\ . \tag{6}$$

This motivates the development of a system model that allows us to exploit the more tractable service representation in Eq. (6). In this alternative model, arrivals, departures, and service reside in a different domain, where we can work directly with the distribution functions of the fading channel gain and the corresponding SNR at the receiver.

C. Network Model in the SNR Domain

We now proceed by mapping the network model from Fig. 1 into a transfer domain, which we refer to as *SNR domain*. We seek to derive performance bounds in the transfer domain, and then map the results to the bit domain to obtain network-layer bounds for backlog and delays. The relationship of the network models in bit domain and SNR domain is illustrated in Fig. 2.

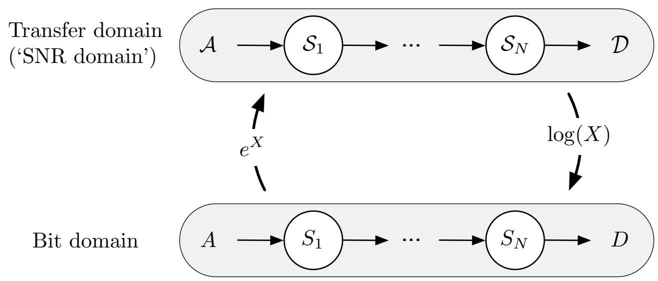

Fig. 2. Transfer domain of network model.

In the previous subsection, we constructed the service process for a wireless link in the SNR domain in Eq. (6) as

$$\mathcal{S}(\tau,t) = e^{S(\tau,t)} .$$

By analogy, we describe the arrivals and departures in the SNR domain respectively by

$$\mathcal{A}(\tau,t) \stackrel{\triangle}{=} e^{A(\tau,t)} \quad \text{and} \quad \mathcal{D}(\tau,t) \stackrel{\triangle}{=} e^{D(\tau,t)} .$$

Throughout this paper, we use calligraphic letters to represent processes that characterize traffic or service as a function of the instantaneous SNR in the sense of Eq. (6). Due to the monotonicity of the exponential function, $\mathcal{D}(0,t)$ and $\mathcal{A}(0,t)$ are increasing in t, and satisfy the causality property $\mathcal{D}(0,t) \leq \mathcal{A}(0,t)$. The backlog process is then described by

$$\mathcal{B}(t) \stackrel{\triangle}{=} e^{B(t)} = \mathcal{A}(0,t)/\mathcal{D}(0,t) .$$

The transformation does not affect time. Therefore,

$$\mathcal{W}(t) \stackrel{\triangle}{=} W(t) = \inf\{u \geq 0 : \mathcal{A}(0,t) \leq \mathcal{D}(0,t+u)\} . \quad (7)$$

To interpret these processes in the transfer domain, let $\gamma_{a,i} = g^{-1}(e^{a_i})$ be the instantaneous channel SNR required to transmit a_i in a single time slot, assuming transmission at the rate of the capacity limit. The arrival process in the SNR domain can then be expressed in terms of these variables as

$$\mathcal{A}(\tau,t) = \prod_{i=\tau}^{t-1} g(\gamma_{a,i}) . \quad (8)$$

Here, we are treating channel quality expressed in terms of the instantaneous SNR as a commodity. An arrival in a time unit represents a workload, where $\gamma_{a,i}$ expresses the amount of resources that will be consumed by the workload. The backlog can similarly be expressed in terms of the instantaneous SNR

$$\mathcal{B}(t) = \prod_{i=t}^{t+\tau_B-1} g(\gamma_i) ,$$

with the interpretation that a node with backlog $B(t)$ at time t requires full use of the channel capacity for τ_B time units to clear the backlog.

Most importantly, the concept of the dynamic server translates to the SNR domain. In a network system, the service process in the bit domain satisfies Eq. (2) if and only if the process in the SNR domain satisfies

$$\mathcal{D}(0,t) \geq \inf_{0\leq u\leq t}\{\mathcal{A}(0,u) \cdot \mathcal{S}(u,t)\} . \quad (9)$$

We refer to a network element that satisfies Eq. (9) for any sample path as *dynamic SNR server*. In this general setting, we do not require that $\mathcal{S}$ takes the form in Eq. (6), in particular, $\mathcal{S}(\tau,t)$ does not have to be equal to $\mathcal{S}(\tau,u)\cdot\mathcal{S}(u,t)$.

Traffic aggregation in the SNR domain is expressed in terms of a product. When M flows have arrivals at a node with arrival processes denoted by $A_k, k = 1,\dots,K$, then the total arrival, A_{agg}, are given for any $0 \leq \tau \leq t$ by $A_{\text{agg}}(\tau,t) = \sum_{k=1}^{K} A_k(\tau,t)$. Since A_{agg} is a sum of random processes it expresses statistical multiplexing gain. If we let $\mathcal{A}_k$ and $\mathcal{A}_{\text{agg}}$ denote the corresponding processes in the SNR domain, we see that

$$\mathcal{A}_{\text{agg}}(\tau,t) = \prod_{k=1}^{K} \mathcal{A}_k(\tau,t) .$$

With the above definitions, the usual network description by a $(\min,+)$ dioid algebra in the bit domain can be expressed in the SNR domain by a dioid algebra on $\mathcal{F}$ where the second operator is a multiplication. This enables the development of the $(\min,\times)$ network calculus in Sec. IV. We observe that the exponential function defines a one-to-one correspondence between arrival and departure processes in the bit and SNR domains. The physical arrival, departure, service, and backlog processes can be recovered from their counterparts in the SNR domain by taking a logarithm (see Fig. 2).

IV. Stochastic $(\min,\times)$ Network Calculus

This section contains our main contribution: the derivation of statistical end-to-end performance bounds for a network where service is expressed in terms of fading distributions residing in the SNR domain. A key characteristic of the approach is that it does not require secondary (network-layer) models of fading distributions when expressed in the SNR domain, that is, bounds are expressed in terms of the fading parameters of the channel model.

By an *SNR process* we mean a bivariate process $\mathcal{X}(\tau,t)$ taking values in $\mathbb{R}^+$ that is increasing in the second argument, with $\mathcal{X}(t,t) = 1$ for all t. The space of SNR processes will be denoted by $\mathcal{F}^+$. For any pair of SNR processes $\mathcal{X}(\tau,t)$ and $\mathcal{Y}(\tau,t)$, set

$$\mathcal{X} \otimes \mathcal{Y}(\tau,t) \stackrel{\triangle}{=} \inf_{\tau\leq u\leq t} \left\{\mathcal{X}(\tau,u)\cdot\mathcal{Y}(u,t)\right\} , \quad (10)$$

and

$$\mathcal{X} \oslash \mathcal{Y}(\tau,t) \stackrel{\triangle}{=} \sup_{u\leq\tau} \left\{\frac{\mathcal{X}(u,t)}{\mathcal{Y}(u,\tau)}\right\} . \quad (11)$$

We refer to '$\otimes$' and '$\oslash$' as the $(\min,\times)$ convolution and $(\min,\times)$ deconvolution operators, respectively.

The arrival, departure, and service processes constructed in the previous section are SNR processes. With the $(\min,\times)$ convolution, we can express the defining property of a dynamic SNR server from Eq. (9) as

$$\mathcal{D}(0,t) \geq \mathcal{A} \otimes \mathcal{S}(0,t) \quad (12)$$

for every pair of SNR arrival and departure processes $\mathcal{A}(\tau,t)$ and $\mathcal{D}(\tau,t)$.

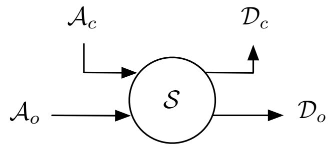

Fig. 3. Single fading channel with cross traffic.

We note that, in fact, for any system description in the bit domain by the $(\mathbb{R} \cup \{+\infty\}, \min, +)$ and the $(\mathcal{F}, \min, *)$ dioid algebras there exists a corresponding description in the SNR domain using $(\mathbb{R}^+ \cup \{+\infty\}, \min, \times)$ and $(\mathcal{F}^+, \min, \otimes)$ dioids.

A. *Wireless Node With Cross Traffic*

Consider a scenario in Fig. 3 where a through flow arriving to a fading channel shares the available bandwidth with other flows. We will refer to the traffic from these other flows as *cross traffic*. We use $\mathcal{A}_o(\tau, t)$ and $\mathcal{A}_c(\tau, t)$ to denote the SNR arrival processes of the through flow and the cross traffic, respectively, and let $\mathcal{D}_o(\tau, t)$ and $\mathcal{D}_c(\tau, t)$ denote the corresponding departure processes. In the SNR domain, cross traffic can be viewed as reducing the channel capacity of the through flow by generating interference.

The following lemma states that, in the SNR domain, the service available to a through flow that experiences cross traffic at a channel can be expressed by a dynamic SNR server.

Lemma 1. *Consider a channel with a through flow and cross traffic as shown in Fig. 3. Assume that the channel provides a dynamic SNR server to the aggregate of through flow and cross traffic, with service process $\mathcal{S}(\tau, t)$, i.e,*

$$\mathcal{D}_o(0, t) \cdot \mathcal{D}_c(0, t) \geq (\mathcal{A}_o \cdot \mathcal{A}_c) \otimes \mathcal{S}(0, t) \,.$$

Then

$$\mathcal{S}_o(\tau, t) = \frac{\mathcal{S}(\tau, t)}{\mathcal{A}_c(\tau, t)}$$

is a dynamic SNR server satisfying for all $t \geq 0$ that

$$\mathcal{D}_o(0, t) \geq \mathcal{A}_o \otimes \mathcal{S}_o(0, t) \,. \tag{13}$$

We refer to the process $\mathcal{S}_o(\tau, t)$ as a "*leftover server*." In light of Lemma 1, it is reasonable to view the cross traffic as interference in the SNR domain, that is, cross traffic reduces the SNR of the through traffic.

Proof: For any sample path, and any $t \geq 0$, we have

$$\mathcal{D}_o(0, t) \cdot \mathcal{D}_c(0, t) \geq \inf_{0 \leq \tau \leq t} \{(\mathcal{A}_o(0, \tau) \cdot \mathcal{A}_c(0, \tau)) \cdot \mathcal{S}(\tau, t)\} \,.$$

Let τ^* be the point where the infimum is assumed. Dividing by $\mathcal{D}_c(\tau, t)$, we obtain

$$\begin{aligned} \mathcal{D}_o(0, t) &\geq \frac{(\mathcal{A}_0 \cdot \mathcal{A}_c) \otimes \mathcal{S}(0, t)}{\mathcal{D}_c(0, t)} \\ &\geq \frac{\mathcal{A}_o(0, \tau^*) \cdot \mathcal{A}_c(0, \tau^*) \cdot \mathcal{S}(\tau^*, t)}{\mathcal{D}_c(0, t)} \Big\} \\ &\geq \Big\{ \mathcal{A}_o(0, \tau^*) \cdot \frac{\mathcal{S}(\tau^*, t)}{\mathcal{A}_c(\tau^*, t)} \Big\} \,, \end{aligned}$$

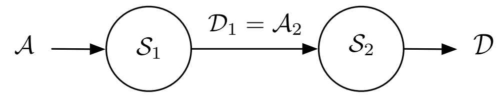

Fig. 4. Tandem of dynamic SNR servers.

where we used that $\mathcal{D}_c(0, t) \leq \mathcal{A}_c(0, t)$ by causality. The lemma follows from the definition of the $(\min, \times)$ convolution. ■

Note that $\mathcal{S}_o(\tau, t)$ need not be monotone in t and may take values below one, i.e., it may not lie in $\mathcal{F}^+$. Monotonicity can be restored by replacing $\mathcal{S}_o$ with a smaller increasing function. $\mathcal{S}_o(\tau, t) \geq 1$ can be ensured when the SNR service of the cross traffic satisfies an upper bound on the departure process, as given in the following corollary.

Corollary 1. *Under the assumptions in Lemma 1, if the service to the cross flows satisfies the upper bound $\mathcal{D}_c(0, t) \leq \mathcal{A}_c \otimes \mathcal{S}(0, t)$, then*

$$\mathcal{D}_o(0, t) \geq \mathcal{A}_o \otimes \max\{1, \mathcal{S}_o\}(0, t) \,. \tag{14}$$

Proof: To prove this claim, assume that $\mathcal{D}_c(0, t) \leq \mathcal{A}_c \otimes \mathcal{S}(0, t)$. Then $\mathcal{D}_c(0, t) \leq \mathcal{A}_c(0, \tau^*) \cdot \mathcal{S}(\tau^*, t)$, and therefore

$$\begin{aligned} \mathcal{D}_o(0, t) &\geq \frac{\mathcal{A}_o(0, \tau^*) \cdot \mathcal{A}_c(0, \tau^*) \cdot \mathcal{S}(\tau^*, t)}{A_c(0, \tau^*) \cdot \mathcal{S}(\tau^*, t)} \\ &= \mathcal{A}_0(\tau^*, t) \,. \end{aligned}$$

Combining this with Lemma 1, we obtain

$$\mathcal{D}_0(0, t) \geq \mathcal{A}_0(0, \tau^*) \cdot \max\{1, \mathcal{S}_o(\tau^*, t)\} \,,$$

proving the corollary. ■

Note that Lemma 1 and Corollary 1 permit descriptions of channel models with memory in the form of time-correlated cross-traffic. When a time-correlated cross traffic process, e.g., a Markov modulated arrival process, is used in the calculation of the leftover service process, the resulting service process in this case is time-correlated as well.

B. *Server Concatenation and Performance Bounds*

The existing network calculus in the bit domain allows for the concatenation of tandem service elements using the $(\min, +)$ convolution (see Sec. I). As an immediate consequence, single node performance bounds are extended to a multi-hop setting. We now establish the corresponding result in the $(\min, \times)$ network calculus. Specifically, the concatenation of dynamic SNR servers is again a dynamic SNR server. We will prove the result for a tandem network of two nodes, as shown in Fig. 4.

Lemma 2. *Let $\mathcal{S}_1(\tau, t)$ and $\mathcal{S}_2(\tau, t)$ be two dynamic SNR servers in tandem as shown in Fig. 4. Then, the service offered by the tandem of nodes is given by the dynamic SNR server $\mathcal{S}_{\mathrm{net}}(\tau, t)$ with*

$$\mathcal{S}_{\mathrm{net}}(\tau, t) = \mathcal{S}_1 \otimes \mathcal{S}_2(\tau, t) \,.$$

Proof: Using Eq. (9), the departure process $\mathcal{D}(0,t)$ can be written as

$$\begin{aligned}\mathcal{D}(0,t) &\geq \inf_{0\leq u\leq t}\{\mathcal{A}_2(0,u)\cdot\mathcal{S}_2(u,t)\}\\ &\geq \inf_{0\leq u\leq t}\Big\{\inf_{0\leq\tau\leq u}\{\mathcal{A}(0,\tau)\cdot\mathcal{S}_1(\tau,u)\}\cdot\mathcal{S}_2(u,t)\Big\}\\ &= \inf_{0\leq\tau\leq t}\Big\{\mathcal{A}(0,\tau)\cdot\inf_{\tau\leq u\leq t}\{\mathcal{S}_1(\tau,u)\cdot\mathcal{S}_2(u,t)\}\Big\}\\ &= \inf_{0\leq\tau\leq t}\Big\{\mathcal{A}(0,\tau)\cdot(\mathcal{S}_1\otimes\mathcal{S}_2)(\tau,t)\Big\}\,.\end{aligned}$$

∎

The extension to networks with more than two nodes follows by iteratively applying Lemma 2. Hence, the dynamic network SNR server with N nodes in tandem is given by

$$\mathcal{S}_{\rm net}(\tau,t)=\mathcal{S}_1\otimes\mathcal{S}_2\otimes\cdots\otimes\mathcal{S}_N(\tau,t)\ . \tag{15}$$

Performance bounds in the $(\min,\times)$ network calculus are computed with the $(\min,\times)$ deconvolution operator. This is analogous to the role of the $(\min,+)$ deconvolution in the existing $(\min,+)$ network calculus. The bounds are expressed in the following lemma.

Lemma 3. *Given a system with SNR arrival process $\mathcal{A}(\tau,t)$ and dynamic SNR server $\mathcal{S}(\tau,t)$.*

- OUTPUT BURSTINESS. *The SNR departure process is bounded by $\mathcal{D}(\tau,t)\leq\mathcal{A}\oslash\mathcal{S}(\tau,t)$.*
- BACKLOG BOUND. *The SNR backlog process is bounded by $\mathcal{B}(t)\leq\mathcal{A}\oslash\mathcal{S}(t,t)$.*
- DELAY BOUND. *The delay process is bounded by $\mathcal{W}(t)\leq\inf\Big\{d\geq 0:\mathcal{A}\oslash\mathcal{S}(t+d,t)\leq 1\Big\}$.*

Proof: For the output bound, we fix τ and t with $0\leq\tau\leq t$ and derive

$$\begin{aligned}\mathcal{D}(\tau,t) &= \frac{\mathcal{D}(0,t)}{\mathcal{D}(0,\tau)}\leq\sup_{0\leq u\leq\tau}\left\{\frac{\mathcal{A}(0,t)}{\mathcal{A}(0,u)\cdot\mathcal{S}(u,\tau)}\right\}\\ &= \sup_{0\leq u\leq\tau}\left\{\frac{\mathcal{A}(u,t)}{\mathcal{S}(u,\tau)}\right\},\end{aligned}$$

where we used the inequality $\mathcal{D}(0,\tau)\geq\mathcal{A}\otimes\mathcal{S}(0,\tau)$ in the second step.

For any fixed sample path, fix an arbitrary $t\geq 0$. The bound on the backlog is derived by

$$\begin{aligned}\mathcal{B}(t) &= \frac{\mathcal{A}(0,t)}{\mathcal{D}(0,t)}\leq\sup_{0\leq u\leq t}\left\{\frac{\mathcal{A}(0,t)}{\mathcal{A}(0,u)\cdot\mathcal{S}(u,t)}\right\}\\ &= \sup_{0\leq u\leq t}\left\{\frac{\mathcal{A}(u,t)}{\mathcal{S}(u,t)}\right\},\end{aligned}$$

where we used $\mathcal{D}(0,t)\geq\mathcal{A}\otimes\mathcal{S}(0,t)$ in the second step.

By definition of the delay in Eq. (7), a delay bound w satisfies

$$\begin{aligned}\mathcal{W}(t) &= \inf\left\{w\geq 0:\frac{\mathcal{A}(0,t)}{\mathcal{D}(0,t+w)}\leq 1\right\}\\ &\leq \inf\left\{w\geq 0:\sup_{0\leq u\leq t}\left\{\frac{\mathcal{A}(0,t)}{\mathcal{A}(0,u)\cdot\mathcal{S}(u,t+w)}\right\}\leq 1\right\}\\ &= \inf\left\{w\geq 0:\sup_{0\leq u\leq t}\left\{\frac{\mathcal{A}(u,t)}{\mathcal{S}(u,t+w)}\right\}\leq 1\right\}. \end{aligned}\tag{16}$$

where we used the inequality $\mathcal{D}(0,t+w)\geq\mathcal{A}\otimes\mathcal{S}(0,t+w)$ in the second line. ∎

With an algebraic description for network performance bounds in the SNR domain in hand, we now turn to the problem of computing the bounds.

C. The Mellin Transform in the SNR domain

The concise expressions from the previous section for the network service and performance bounds in the SNR domain hide the difficulty of computing these expressions. In fact, all expressions of the $(\min,\times)$ network calculus contain products or quotients of random variables. The *Mellin transform* [11] facilitates such computations, particularly when the arrival and service processes are independent.

The Mellin transform of a nonnegative random variable X is defined by

$$\mathcal{M}_X(s)=E[X^{s-1}]. \tag{17}$$

The Mellin transform of a product of two independent random variables X and Y equals the product of their Mellin transforms [11],

$$\mathcal{M}_{X\cdot Y}(s)=E[(X\cdot Y)^{s-1}]=\mathcal{M}_X(s)\cdot\mathcal{M}_Y(s). \tag{18}$$

Similarly, the Mellin transform of the quotient of independent random variables is given by

$$\mathcal{M}_{X/Y}(s)=E[X^{s-1}]E[Y^{1-s}]=\mathcal{M}_X(s)\cdot\mathcal{M}_Y(2-s), \tag{19}$$

where we used independence to factor the expectation.

We will evaluate the Mellin transform only for real valued s, where it is always well-defined (though it may take the value $+\infty$). For every non-negative random variable X, it holds that $\mathcal{M}_X(1)=1$ and $\frac{d}{ds}\mathcal{M}_X(1)=E[\log X]$. When $s>1$, the Mellin transform is order-preserving, i.e., for any pair of random variables X,Y with $Pr(X>Y)=0$ we have $\mathcal{M}_X(s)\leq\mathcal{M}_Y(s)$ for all s. When $s<1$, the order is reversed. Hence bounds on the distribution of a random variable X generally imply bounds on its Mellin transform.

A more subtle question is how to obtain bounds on the distribution of a random variable from its Mellin transform. Here, the complex inversion formula is not helpful. Instead, we use the moment bound

$$Pr(X\geq a)\leq a^{-s}\mathcal{M}_X(1+s) \tag{20}$$

for all $a>0$ and $s>0$. For bivariate random processes $\mathcal{X}(\tau,t)$, we write $\mathcal{M}_{\mathcal{X}}(s,\tau,t)\stackrel{\triangle}{=}\mathcal{M}_{\mathcal{X}(\tau,t)}(s)$.

We work with the Mellin transform of $(\min,\times)$ convolutions and deconvolutions, which not only involves products and quotients, but also requires to compute infimums and supremums. The exact computation of the Mellin transform for these operations is generally not feasible. We therefore resort to bounds, as specified in the next lemma.

Lemma 4. *Let $\mathcal{X}(\tau,t)$ and $\mathcal{Y}(\tau,t)$ be two independent non-negative bivariate random processes. For $s<1$, the Mellin transform of the $(\min,\times)$ convolution $\mathcal{X}\otimes\mathcal{Y}(\tau,t)$ is bounded by*

$$\mathcal{M}_{\mathcal{X}\otimes\mathcal{Y}}(s,\tau,t)\leq\sum_{u=\tau}^{t}\mathcal{M}_{\mathcal{X}}(s,\tau,u)\cdot\mathcal{M}_{\mathcal{Y}}(s,u,t)\,. \tag{21}$$

For $s>1$, the Mellin transform of the $(\min,\times)$ deconvolution $\mathcal{X}\oslash\mathcal{Y}(\tau,t)$ is bounded by

$$\mathcal{M}_{\mathcal{X}\oslash\mathcal{Y}}(s,\tau,t)\leq\sum_{u=0}^{\tau}\mathcal{M}_{\mathcal{X}}(s,u,t)\cdot\mathcal{M}_{\mathcal{Y}}(2-s,u,\tau)\,. \quad (22)$$

Proof: Note that the function $f(z)=z^{s-1}$ is increasing for $s>1$ and decreasing for $s<1$. For $s<1$, the convolution is estimated by

$$\begin{aligned}\mathcal{M}_{\mathcal{X}\otimes\mathcal{Y}}(s,\tau,t)&=E\Big[\Big(\inf_{\tau\leq u\leq t}\{\mathcal{X}(\tau,u)\cdot\mathcal{Y}(u,t)\}\Big)^{s-1}\Big]\\&=E\Big[\sup_{\tau\leq u\leq t}\Big\{(\mathcal{X}(\tau,u))^{s-1}\cdot(\mathcal{Y}(u,t))^{s-1}\Big\}\Big]\\&\leq\sum_{u=\tau}^{t}E\big[(\mathcal{X}(\tau,u))^{s-1}\big]\cdot E\big[(\mathcal{Y}(u,t))^{s-1}\big]\,.\end{aligned}$$

In the last step, we have used the non-negativity of $\mathcal{X}$ and $\mathcal{Y}$ and the union bound to replace the supremum with a sum, and their independence to evaluate the expectation of the products. Eq. (21) follows by inserting the definition of the Mellin transform. The deconvolution is similarly estimated for $s>1$ by

$$\begin{aligned}\mathcal{M}_{\mathcal{X}\oslash\mathcal{Y}}(s,\tau,t)&=E\Big[\Big(\sup_{u\leq\tau}\{\mathcal{X}(u,t)/\mathcal{Y}(u,\tau)\}\Big)^{s-1}\Big]\\&=E\Big[\sup_{0\leq u\leq\tau}\Big\{\big(\mathcal{X}(u,t)\big)^{s-1}\cdot\big(\mathcal{Y}(u,\tau)\big)^{1-s}\Big\}\Big]\\&\leq\sum_{u=0}^{\tau}E\Big[\big(\mathcal{X}(u,t)\big)^{s-1}\cdot\big(\mathcal{Y}(u,\tau)\big)^{1-s}\Big]\,, \quad (23)\end{aligned}$$

and Eq. (22) follows from the independence assumption and the definition of the Mellin transform. ■

As a remark, in the lemma we assumed that the arrival and service processes are independent. This is a convenient and often justifiable assumption, however, it limits the applicability of the obtained results to wireless systems that exhibit unmitigated co-channel interference which introduces dependence between service processes in a multi-hop setting. For the analysis of processes that are dependent weaker bounds can be obtained by applying the Hölder inequality. For example, if the two processes $\mathcal{X}(\tau,t)$ and $\mathcal{Y}(\tau,t)$ in Eq. (23) are dependent, then we can apply the Hölder inequality to the last step and bound their $(\min,\times)$ deconvolution by

$$\begin{aligned}\mathcal{M}_{\mathcal{X}\oslash\mathcal{Y}}(s,\tau,t)&\leq\sum_{u=0}^{\tau}\Big(E\Big[(\mathcal{X}(u,t))^{(s-1)p}\Big]\Big)^{1/p}\\&\qquad\cdot\Big(E\Big[(\mathcal{Y}(u,\tau))^{(1-s)q}\Big]\Big)^{1/q}\\&=\sum_{u=0}^{\tau}\Big(\mathcal{M}_{\mathcal{X}}(1-p+sp,u,t)\Big)^{1/p}\cdot\Big(\mathcal{M}_{\mathcal{Y}}(1+q-sq,u,\tau)\Big)^{1/q}\end{aligned}$$

for any $p>1$ and $1/p+1/q=1$. We used the positivity of $\mathcal{X}$ and $\mathcal{Y}$ and applied the Hölder inequality in the first step and the definition of the Mellin transform in the second step. This bound can be optimized over the choice of p.

D. Performance Bounds for the Bit Domain

We next obtain network-level performance bounds for the bit domain. This involves a transformation from the SNR domain to the bit domain via the relationship in Fig. 2.

Theorem 1. *Given a system where arrivals are described by a bivariate process $A(\tau,t)$, and the available service is given by a dynamic server $S(\tau,t)$. Let $\mathcal{A}(\tau,t)$ and $\mathcal{S}(\tau,t)$ be the corresponding SNR processes. Fix $\varepsilon>0$ and define, for $s>0$,*

$$\mathsf{M}(s,\tau,t)\triangleq\sum_{u=0}^{\min(\tau,t)}\mathcal{M}_{\mathcal{A}}(1+s,u,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,u,\tau)\,.$$

Then, we have the following probabilistic performance bounds.

- OUTPUT BURSTINESS: $Pr\big(D(\tau,t)>d^{\varepsilon}\big)\leq\varepsilon$, *where*

$$d^{\varepsilon}(\tau,t)=\inf_{s>0}\Big\{\frac{1}{s}\big(\log\mathsf{M}(s,\tau,t)-\log\varepsilon\big)\Big\}\,;$$

- BACKLOG: $Pr\big(B(t)>b^{\varepsilon}\big)\leq\varepsilon$, *where*

$$b^{\varepsilon}=\inf_{s>0}\Big\{\frac{1}{s}\big(\log\mathsf{M}(s,t,t)-\log\varepsilon\big)\Big\}\,;$$

- DELAY: $Pr\big(W(t)>w^{\varepsilon}\big)\leq\varepsilon$, *where w^{ε} is the smallest number satisfying*

$$\inf_{s>0}\Big\{\mathsf{M}(s,t+w^{\varepsilon},t)\Big\}\leq\varepsilon\,.$$

If stability of the system is not assured, the bounds in the theorem may not be finite or grow over time. We address stability conditions and the tightness of the bounds in the next subsection.

Proof: Lemma 3 defines the three performance bounds in terms of the $(\min,\times)$ deconvolution of $\mathcal{A}$ and $\mathcal{S}$. For the bound on the distribution of the output burstiness, we start from the inequality $\mathcal{D}(\tau,t)\leq\mathcal{A}\oslash\mathcal{S}(\tau,t)$. It follows from the moment bound and Lemma 4 that, for any choice of $d>0$ and all $s>0$

$$\begin{aligned}Pr(D(\tau,t)>d)&=Pr(\mathcal{D}(\tau,t)>e^{d})\\&\leq Pr(\mathcal{A}\oslash\mathcal{S}(\tau,t)>e^{d})\\&\leq(e^{d})^{-s}\mathcal{M}_{\mathcal{A}\oslash\mathcal{S}}(1+s,\tau,t)\\&\leq e^{-sd}\mathsf{M}(s,\tau,t)\,.\end{aligned}$$

To obtain the claim, we set the right hand side equal to ε, solve for d, and optimize over the value of $s>0$ to obtain $d^{\varepsilon}(\tau,t)$. The proof of the backlog bound proceeds in the same way, starting from the inequality $\mathcal{B}(t)\leq\mathcal{A}\oslash\mathcal{D}(t,t)$, resulting in

$$Pr(B(t)>b)\leq e^{-sb}\mathsf{M}(s,t,t)\,. \quad (24)$$

The delay bound is slightly more subtle. Fix $t\geq0$. Using Lemma 3 and the moment bound with $a=1$, we obtain that

$$\begin{aligned}Pr(\mathcal{W}(t)>w)&\leq Pr(\mathcal{A}\oslash\mathcal{S}(t+w,t)>1)\\&\leq\mathcal{M}_{\mathcal{A}\oslash\mathcal{S}}(1+s,t+w,t)\end{aligned}$$

for every $s>0$. By Lemma 4, the Mellin transform $\mathcal{M}_{\mathcal{A}\oslash\mathcal{S}}(1+s,t+w,t)$ satisfies a bound that agrees with the function $\mathsf{M}(s,t+w,t)$, except that the upper limit in the summation that defines $\mathsf{M}(s,t+w,t)$ would have to be replaced by $\tau = t+w$. To obtain the sharper estimate from the claim, we use instead Eq. (16) from the proof of Lemma 3. The resulting bound is that

$$\mathcal{Z}(t) \stackrel{\triangle}{=} \sup_{0\le u\le t}\left\{\frac{\mathcal{A}(u,t)}{\mathcal{S}(u,t+w)}\right\}$$

satisfies

$$\begin{aligned} Pr(\mathcal{W}(t) > w) &\le Pr(\mathcal{Z}(t) > 1) \\ &\le \mathcal{M}_{\mathcal{Z}(t)}(1+s) \\ &\le \mathsf{M}(s, t+w, t)\,, \end{aligned} \tag{25}$$

where we have used that the supremum in the definition of $\mathcal{Z}$ extends only up to $u = t$, and then repeated the proof of Eq. (22). The claim follows by optimizing over s. ■

Corresponding bounds as in Theorem 1 can be obtained using the $(\min,+)$ algebra and the network calculus with moment-generating functions [12]. The significance of Theorem 1 is that it permits the application of network calculus, where traffic is characterized in the bit domain, and service is naturally expressed in the SNR domain. This will become evident in Sec. V, where the theorem gives us concise bounds for delays and backlog in multi-hop networks with Rayleigh fading channels.

E. Asymptotic Tightness of the Bounds

In this section, we show that the upper bounds in Theorem 1 have an exponential rate of decay, and that the rate of decay cannot be improved without adding assumptions. We show the derivations for the backlog. We require that arrival and service processes are stationary. We assume that the average service rate exceeds the average arrival rate, i.e.,

$$\frac{1}{t}E[S(0,t)] > \frac{1}{t}E[A(0,t)]\,. \tag{26}$$

We will see that this inequality, which, by stationarity, does not depend on t, is the stability condition of the system. It guarantees that the backlog process $B(t)$ is stochastically increasing in t, and converges in distribution to the steady-state backlog process $B_\infty = \lim_{t\to\infty} B(t)$ [26, Lemma 2].

We start with a corollary to Theorem 1 that provides an exponential decay rate for the backlog bound. Define

$$\Lambda(s) \stackrel{\triangle}{=} \frac{1}{t}\log\left(\mathcal{M}_{\mathcal{A}}(1+s,0,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,0,t)\right)\,. \tag{27}$$

The function Λ is related to expressions for the effective bandwidth [20] or effective service [39]. It is convex because it is defined as a limit of convex functions. Since $\Lambda(0) = 0$ and

$$\frac{d}{ds}\Lambda(0) = \lim_{t\to\infty}\frac{1}{t}(E[A(0,t)] - E[S(0,t)]) < 0\,,$$

$\Lambda(s)$ either changes sign exactly once from negative to positive, or $\Lambda(s) < 0$ for all $s > 0$. Let s^* denote the point where the switch occurs, with $s^* = \infty$ when $\Lambda(s)$ remains negative. If $s^* < \infty$ and Λ is differentiable, we note for later use that the convexity of Λ forces $\frac{d}{ds}\Lambda(s^*) > 0$.

Corollary 2. *For each s with $0 < s < s^*$ there is a constant $C(s)$ such that the backlog bound*

$$Pr(B(t) > x) \le C(s)e^{-sx} \tag{28}$$

holds for all $t \ge 0$.

The constant $C(s)$ in the corollary depends on the arrival and service processes, but not on t.

Proof: Consider the backlog bound in Eq. (24). We want to show that

$$C(s) = \sup_{t\ge0}\mathsf{M}(s,t,t) < \infty\,.$$

The constant can be evaluated as

$$\begin{aligned} C(s) &= \sup_{t\ge0}\sum_{u=0}^{t}\mathcal{M}_{\mathcal{A}}(1+s,u,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,u,t) \\ &= \sum_{u=0}^{\infty}\mathcal{M}_{\mathcal{A}}(1+s,0,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,0,t)\,. \end{aligned}$$

We have used stationarity to replace the time interval $[u,t)$ with $[0,t-u)$, changed the variable from u to $t-u$, and then taken $t\to\infty$. For $0 < s < s^*$, we have $\Lambda(s) < 0$. By Eq. (27) there exists an $\varepsilon > 0$ and a time t_ε such that

$$\frac{1}{t}\log\big(\mathcal{M}_{\mathcal{A}}(1+s,0,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,0,t)\big) \le -\varepsilon$$

for all $t \ge t_\varepsilon$. It follows that

$$\mathcal{M}_{\mathcal{A}}(1+s,0,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,0,t) \le e^{-\varepsilon t}$$

for $t \ge t_\varepsilon$. Therefore, the sum that determines $C(s)$ converges for all $0 < s < s^*$. ■

Since $C(s)$ does not depend on t, Eq. (28) holds also for the steady-state backlog process B_∞. Solving for s and then taking the limit $s \to s^*$, we see that

$$\lim_{x\to\infty}\frac{1}{x}\log Pr(B_\infty \ge x) \le -s^*\,.$$

We now show that this bound on the decay rate is tight in the sense that it cannot be improved without adding further assumptions to Theorem 1. To this end, we consider the special case of a dynamic server $\mathcal{S}(\tau,t)$ that satisfies Eq. (9) with equality, i.e., $\mathcal{D}(0,t) = \mathcal{A}\otimes\mathcal{S}(0,t)$. An example of this is the fading channel service model described in Section III.A.

We also need the following technical conditions. The function $\Lambda(s)$ should be defined and differentiable on some interval $[0, s_{\max})$, and its derivative should be unbounded from above. Moreover, we assume that $\Lambda(s)$ changes sign, so that $s^* < \infty$. Under these assumptions, we will obtain an exponential lower bound on B_∞ for every $s > s^*$.

In a similar fashion as in [6, Chapter 9] for showing that the effective bandwidth offers a lower bound on resource requirements, we will apply large deviations theory, specifically, the Gärtner-Ellis theorem. In the context of the SNR domain and with our notation, the theorem as given in [32, Theorem D.8] takes the following form:

Lemma 5 (Gärtner-Ellis)**.** *Let $\{\mathcal{X}_t\}_{t\geq 1}$ be a sequence of nonnegative random variables, and let $\mathcal{M}_{\mathcal{X}}(s,t)$ denote their Mellin transform. Assume that*

$$\Lambda(s) = \lim_{t\to\infty} \frac{1}{t} \log \mathcal{M}_{\mathcal{X}}(1+s,t)\,,$$

exists for $s \in [0, s_{\max})$, and defines a differentiable function whose derivative is unbounded from above. Then

$$\lim_{t\to\infty} \frac{1}{t} \log Pr(\mathcal{X}_t > e^{\beta t}) = -I(\beta)\,,$$

where the rate function I is given by the Legendre transform

$$I(\beta) \stackrel{\triangle}{=} \sup_s \{\beta s - \Lambda(s)\}\,. \tag{29}$$

We will apply Lemma 5 to the process $\mathcal{X}_t = \mathcal{A}(0,t)/\mathcal{S}(0,t)$. With our assumptions, $\Lambda(s)$ from Eq. (27) satisfies the conditions needed for the Gärtner-Ellis theorem.

Theorem 2. *Under the assumptions stated before Lemma 5,*

$$\liminf_{x\to\infty} \frac{1}{x} \log Pr(B_\infty \geq x) \geq -s^*\,, \tag{30}$$

where s^* is the unique positive solution of $\Lambda(s) = 0$.

The theorem implies that there exists for each $s > s^*$ a constant $C(s)$ such that

$$Pr(B_\infty > x) \geq C(s) e^{-sx}$$

for all $x > 0$. Together with Corollary 2, it extends [6, Theorem 9.1.1] to variable rate servers.

Proof: We may assume that $s^* < \infty$ (otherwise, there is nothing to show). We start from the observation that $\mathcal{D}(0,t) \geq \mathcal{A} \otimes \mathcal{S}(0,t)$ implies

$$\mathcal{B}(t) = \sup_{0\leq u\leq t} \left\{ \frac{\mathcal{A}(u,t)}{\mathcal{S}(u,t)} \right\} \geq \frac{\mathcal{A}(0,t)}{\mathcal{S}(0,t)}\,.$$

It follows that

$$\begin{aligned} Pr(B_\infty \geq x) &\geq Pr(B(t) \geq x) \\ &\geq Pr\Big(\mathcal{A}(0,t)/\mathcal{S}(0,t) \geq e^x\Big) \end{aligned}$$

for every $x \geq 0$ and all $t \geq 0$. Let $\beta > 0$ be arbitrary. Setting $x = \beta t$ and taking $t \to \infty$, we have

$$\begin{aligned} \liminf_{x\to\infty} & \frac{1}{x} \log Pr(B_\infty \geq x) \\ &\geq \frac{1}{\beta} \lim_{t\to\infty} \frac{1}{t} \log Pr(\mathcal{A}(0,t)/\mathcal{S}(0,t) \geq e^{\beta t}) \\ &= -\frac{1}{\beta} I(\beta)\,, \end{aligned} \tag{31}$$

where Lemma 5 is used in the last step. Finally, we choose $\beta^* = \frac{d}{ds}\Lambda(s^*) > 0$. For $\beta = \beta^*$, the supremum in Eq. (29) is assumed at $s = s^*$. The proof is completed by inserting the value $I(\beta^*) = \beta^* s^*$ into Eq. (31). ■

V. Network Performance of Rayleigh Channels

We now apply the techniques developed in the two previous sections to a network of Rayleigh channels. Consider the dynamic SNR server description for a Rayleigh fading channel, as constructed in Sec. III.B. We use Eq. (6), with the function $g(\gamma)$ given by

$$g(\gamma) = 1 + \gamma = 1 + \bar{\gamma}|h|^2, \tag{32}$$

where $\bar{\gamma}$ is the average SNR of the channel and $|h|$ is the fading gain. For Rayleigh fading, $|h|$ is a Rayleigh random variable with probability density $f(x) = 2xe^{-x^2}$. In a physical system, $\bar{\gamma} = \bar{P}_r/\sigma^2$, where $\bar{P}_r$ and σ^2 are the received signal power and the (additive white Gaussian) noise power at the receiver, respectively. Then, $|h|^2$ is exponentially distributed, and the Mellin transform of $g(\gamma)$ is given by

$$\mathcal{M}_{g(\gamma)}(s) = e^{1/\bar{\gamma}} \bar{\gamma}^{s-1} \Gamma(s, \bar{\gamma}^{-1})\,, \tag{33}$$

where $\Gamma(s,y) = \int_y^\infty x^{s-1} e^{-x}\,dx$ is the upper incomplete Gamma function. Using the assumption that the γ_i are independent and identically distributed, we obtain for the Mellin transform of the dynamic server

$$\mathcal{M}_{\mathcal{S}}(s,\tau,t) = \Big(e^{1/\bar{\gamma}} \bar{\gamma}^{s-1} \Gamma(s, \bar{\gamma}^{-1})\Big)^{t-\tau}. \tag{34}$$

A. Arrival Model

For the arrival process, we use a characterization due to Chang [6], referred to as $(\sigma(s), \rho(s))$-bounded arrivals, where the moment-generating function of the cumulative arrival process in the bit domain is bounded by

$$\frac{1}{s} \log E[e^{sA(\tau,t)}] \leq \rho(s)\cdot(t-\tau) + \sigma(s)$$

for some $s > 0$. In general, $\rho(s)$ and $\sigma(s)$ are nonnegative increasing functions of s that may become infinite when s is large. This characterization can be viewed as a probabilistic extension of a traffic flow that is deterministically regulated by a token bucket with rate ρ and burst size σ. It describes a large traffic class that is well suited to express traffic with burstiness, including bursty traffic with memory, such as Markov modulated On-Off traffic [6].

This class of arrival processes can be equivalently characterized in the SNR domain as a bound on the Mellin transform of the SNR arrival process as follows

$$\mathcal{M}_{\mathcal{A}}(s,\tau,t) \leq e^{(s-1)\cdot(\rho(s-1)\cdot(t-\tau)+\sigma(s-1))}\,, \tag{35}$$

for some $s > 1$. This traffic class is also referred to as $(\sigma(s), \rho(s))$-bounded arrivals.

To describe a Markov-modulated arrival process characterized as $(\sigma(s), \rho(s))$-bounded arrivals, we begin with the log MGF of Markov modulated arrival process, which is given by [6]

$$\lim_{t\to\infty} \frac{1}{st} \log E\big[e^{sA(t)}\big] = \frac{1}{s} \log sp(\mathbf{\Phi}(s)\mathbf{P})\,,$$

where, $\mathbf{\Phi}(s) = \mathrm{diag}\,(\phi_1(s), \ldots, \phi_M(s))$, $\phi_i(s) = E[e^{sr_i}]$, r_i is the transmission rate when in state i, M is the cardinality of the state space, $\mathbf{P}$ is the state transition matrix with p_{ij}

the element in row i and column j, and $sp(C)$ is the spectral radius of matrix C, i.e., the largest eigenvalue of the matrix.

In the numerical examples, we will use a Markov modulated On-Off process, which has two states ($M = 2$), where index 1 indicates the Off state ($r_1 = 0$) and index 2 indicates the On state ($r_2 = r$), and where rate r is referred to as the peak rate. The transition probabilities are $p_{12} = \alpha$ and $p_{21} = \beta$. Then using $(\sigma(s), \rho(s))$ arrival description, we have $\sigma(s) = 0$ and $\rho(s) = \frac{1}{s} \log L(s, r, \alpha, \beta)$, where,

$$L(s, r, \alpha, \beta) \stackrel{\triangle}{=} \frac{e^{sr}(1-\beta)+(1-\alpha)+\sqrt{(e^{sr}(1-\beta)-(1-\alpha))^2+4e^{sr}(1-\alpha-\beta)}}{2}.$$

Since, for real-valued s, Mellin transform of an SNR process is related to the MGF of its bit domain equivalent by

$$E\Big[e^{sX(u,t)}\Big] = E\big[(\mathcal{X}(u,t))^s\big] = \mathcal{M}_{\mathcal{X}}(s+1, u, t)\,,$$

we can compute the Mellin transform for SNR arrivals from the MGF of the corresponding processes in bit domain. We obtain

$$\mathcal{M}_{\mathcal{A}}(s, \tau, t) \le e^{(s-1)\rho(s-1)(t-\tau)} = \big(L(s-1, r, \alpha, \beta)\big)^{t-\tau}. \tag{36}$$

B. Network Service Description

Consider a cascade of fading channels, as in Fig. 1, which is traversed by a through flow. Each fading channel is described by a dynamic SNR server $\mathcal{S}$ satisfying Eq. (6), and the cross traffic at each channel is $\mathcal{A}_c$ satisfying Eq. (35) with parameters $\sigma_c(s)$ and $\rho_c(s)$. Assume that arrivals from through flow and cross traffic, as well as the service processes at each channel are independent. Then we can compute a bound for the Mellin transform of the SNR service process for the through flow in the entire cascade.

Lemma 6. *Given a cascade of N fading channels with cross traffic described above. Let $\mathcal{S}_{o,\text{net}}$ denote the dynamic SNR server that describes the service of the through flow in the cascade. The Mellin transform of $\mathcal{S}_{o,\text{net}}(\tau, t)$ satisfies for $s < 1$ that*

$$\mathcal{M}_{\mathcal{S}_{o,\text{net}}}(s, \tau, t) \le e^{(1-s)\cdot N\sigma_c(1-s)} \binom{N-1+t-\tau}{t-\tau} \cdot \big(\mathcal{M}_{g(\gamma)}(s) e^{(1-s)\cdot\rho_c(1-s)}\big)^{t-\tau}. \tag{37}$$

Proof: For a single channel ($N = 1$), we obtain with Lemma 1 that for $0 \le \tau \le t$

$$\begin{aligned}\mathcal{M}_{\mathcal{S}_o}(s, \tau, t) &= \mathcal{M}_{\mathcal{S}/\mathcal{A}_c}(s, \tau, t)\\ &= \mathcal{M}_{\mathcal{S}}(s, \tau, t) \cdot \mathcal{M}_{\mathcal{A}_c}(2-s, \tau, t)\\ &\le e^{(1-s)\cdot\sigma_c(1-s)} \cdot \big(\mathcal{M}_{g(\gamma)}(s) \cdot e^{(1-s)\cdot\rho_c(1-s)}\big)^{t-\tau},\end{aligned}$$

where we have used that $2-s > 1$ for $s < 1$. When the service of the through flow at the n-th channel is denoted by $\mathcal{S}_{o,n}$, by Lemma 2, the service of the cascade of channels is given by the $(\min, \times)$ convolution $\mathcal{S}_{o,\text{net}}(\tau, t) = \mathcal{S}_{o,1} \otimes \ldots \otimes \mathcal{S}_{o,N}(\tau, t)$. We use Lemma 4 to bound its Mellin transform by

$$\mathcal{M}_{\mathcal{S}_{\text{net}}}(s, \tau, t) \le \sum_{u_1,\ldots,u_{N-1}} \prod_{n=1}^{N} \mathcal{M}_{\mathcal{S}_{o,n}}(s, u_{n-1}, u_n)\,,$$

where the sum runs over all sequences $u_0 \le u_1 \le \cdots \le u_N$ with $u_0 = \tau$ and $u_N = t$. Each product appearing in the equation evaluates to the same term

$$\begin{aligned}\prod_{n=1}^{N} \mathcal{M}_{\mathcal{S}_{o,n}}(s, u_{n-1}, u_n) &= e^{(1-s)\cdot N\sigma_c(1-s)}\\ &\quad\cdot \prod_{n=1}^{N} \big(\mathcal{M}_{g(\gamma)}(s) \cdot e^{(1-s)\cdot(\rho_c(1-s)}\big)^{u_n-u_{n-1}}\\ &= e^{(1-s)\cdot N\sigma_c(1-s)} \big(\mathcal{M}_{g(\gamma)}(s) \cdot e^{(1-s)\cdot(\rho_c(1-s)}\big)^{t-\tau}.\end{aligned}$$

Then we collect terms with the binomial identity

$$\sum_{u_{N-1}=\tau}^{t} \cdots \sum_{u_1=\tau}^{u_2} 1 = \binom{N-1+t-\tau}{t-\tau}\,,$$

and the claim follows. ■

For a cascade of fading channels with no cross traffic, we have $\sigma_c(s) = \rho_c(s) = 0$ and Eq. (37) reduces to

$$\mathcal{M}_{\mathcal{S}_{\text{net}}}(s, \tau, t) \le \binom{N-1+t-\tau}{t-\tau} \cdot \big(\mathcal{M}_{g(\gamma)}(s)\big)^{t-\tau}.$$

C. Performance Bounds for Rayleigh Fading Channels

Now we consider the performance of $(\sigma(s), \rho(s))$-bounded through traffic with parameters $\sigma_o(s)$ and $\rho_o(s)$ in a cascade of Rayleigh fading channel with cross traffic, where cross traffic at each channel is $(\sigma(s), \rho(s))$-bounded with parameters $\sigma_c(s)$ and $\rho_c(s)$. The Mellin transform for the Rayleigh fading channel is given by Eq. (33), and those for through and cross traffic by Eq. (35). We compute end-to-end performance bounds using Theorem 1. Using Lemma 6, we compute the function $\mathsf{M}(s, \tau, t)$ for $0 \le \tau \le t$ as

$$\mathsf{M}_{o,\text{net}}(s, \tau, t) \le e^{s\cdot(\rho_o(s)(t-\tau)+\sigma_o(s)+N\sigma_c(s))} \sum_{u=[\tau-t]_+}^{\infty} \binom{N-1+u}{u} \cdot \big(V_o(s)\big)^u, \tag{38}$$

where $[\tau - t]_+$ is the maximum of $\tau - t$ and 0 and

$$V_o(s) = e^{s\cdot(\rho_o(s)+\rho_c(s))} e^{1/\bar{\gamma}} \bar{\gamma}^{-s} \Gamma(1-s, \bar{\gamma}^{-1}).$$

The sum in Eq. (38) converges when $V_o(s) < 1$. This is ensured for some value of s when the stability condition in Eq. (26) is satisfied.

When $t \ge \tau$, e.g., when computing output burstiness or backlog bounds, we have

$$\mathsf{M}_{o,\text{net}}(s, \tau, t) \le \frac{e^{s\cdot(\rho_o(s)(t-\tau)+\sigma_o(s)+N\sigma_c(s))}}{(1-V_o(s))^N},$$

where we applied the combinatorial identity from Lemma 7 in the Appendix. Then, using Theorem 1, the output burstiness of the through flow at the network egress for a violation probability ε is given by

$$d^{\varepsilon}_{o,\mathrm{net}}(\tau,t) \le \inf_{s>0} \Bigg\{ \sigma_o(s) + N\sigma_c(s) + \rho_o(s)(t-\tau) - \frac{1}{s}\Big[N \log\big(1 - V_o(s)\big) + \log \varepsilon\Big] \Bigg\} \qquad (39)$$

and the end-to-end backlog of the through flow is bounded by

$$b^{\varepsilon}_{o,\mathrm{net}}(t) \le \inf_{s>0} \Bigg\{ \sigma_o(s) + N\sigma_c(s) - \frac{1}{s}\Big[N \log\big(1 - V_o(s)\big) + \log \varepsilon\Big] \Bigg\}. \qquad (40)$$

For the delay bound, we estimate for $w \ge 0$ that

$$\begin{aligned} &\mathsf{M}_{\mathrm{net}}(s, t+w, t) \\ &\quad\le\ e^{s(-\rho_o(s)w + \sigma_o(s) + N\sigma_c(s))} \cdot \sum_{u=w}^{\infty} \binom{N-1+u}{u} (V_o(s))^u \\ &\quad\le \inf_{s>0} \Bigg\{ \frac{e^{s(-\rho_o(s)w + \sigma_o(s) + N\sigma_c(s))}}{(1 - V_o(s))^N} \cdot \min\big\{1, (V_o(s))^w (w+1)^{N-1}\big\} \Bigg\}. \qquad (41) \end{aligned}$$

Here, the first term in the minimum is obtained by extending the summation in Eq. (38) down to $u = 0$ and applying Lemma 7. The second term results from Lemma 8 in the Appendix.

The delay bound w^{ε} is determined according to Theorem 1 by setting the right hand side of Eq. (41) equal to ε, solving for w, and minimizing over s. Because of the complexity of the bound in Eq. (41), the last two steps must be performed numerically. The performance bounds for networks without cross traffic are obtained by inserting the values $\sigma_c(s) = 0$ and $\rho_c(s) = 0$ in the performance bounds expressions derived above.

It is apparent that the complexity of computing end-to-end bounds is no different than bounds for a single channel. More importantly, we observe that the end-to-end bounds scale linearly in the number of nodes N.

VI. Numerical Examples

In this section, we present numerical results for a cascade of N Rayleigh channels with a transmission bandwidth of $W = 20$ kHz, using the expressions we derived in the previous section. We consider two cases: a network with and without cross traffic. For through traffic in both cases, we use $(\sigma(s), \rho(s))$-bounded arrivals with default values $\sigma(s) = 50$ kb and $\rho(s) = 30$ kbps for all values of s, i.e., the rate and burst size are deterministic and correspond to a traffic flow that is shaped by a token bucket with a given rate and burst size. By choosing a deterministic model for the through traffic, the remaining sources of randomness are those of the channel and the cross traffic. Thus, in examples without cross-traffic, we can study how fading channel variability impacts network performance. In examples with cross-traffic, we can observe the relative impact of random cross traffic and channel conditions on network performance.

For the cross traffic we use a Markov-modulated On-Off traffic model, characterized as $(\sigma(s), \rho(s))$-bounded traffic, where at each node we use identical parameters $\sigma_c(s)$ and $\rho_c(s)$. This represents a scenario composed of a cascade of N uniformly spaced wireless nodes in a static environment, hence $\bar{\gamma}_i = \bar{\gamma}$ for all $i = 1, \ldots, N$, which experience Rayleigh fading, i.e., severe fading with no line–of–sight component.

To evaluate the quality of the derived bounds we also include a simulation of a tandem of queues with the parameters above. Since simulations of large networks become computationally prohibitive, the comparison with simulations uses a scenario with at most 10 nodes. The simulations use a fluid-flow arrival and service model, in a time slotted system with intervals of 1 ms. Simulations are run for 10^{10} time slots for one node, and are increased to 10^{11} time slots for 10 nodes.

Recall that existing performance analyses of fading channels generally rely on Markov channel or other secondary models of the fading channels. Since these models involve additional parameter selections, and the accuracy of the selections with regard to actual channel conditions cannot be determined, we do not attempt a comparison of our numerical results to those of prior analyses.

A. *Performance Bounds Without Cross Traffic*

We eliminate cross traffic by setting $\sigma_c(s) = \rho_c(s) = 0$. For a violation probability of $\varepsilon = 10^{-4}$, in Fig. 5 we show the end-to-end backlog for a cascade of N Rayleigh channels, as a function of the average SNR of each channel. Even though the backlog bounds increase only linearly in the number of nodes, the per-node requirements – at least for the last node of the cascade – must satisfy the end-to-end bounds, since it cannot be assumed that backlog is equally distributed across the nodes. When the SNR of the nodes is sufficiently high, the backlog remains low even for a large number of hops. We observe that the channel becomes saturated for $\bar{\gamma} = 5$ dB. When the number of nodes is small, the backlog increases sharply in the vicinity of $\bar{\gamma} = 5$ dB, but remains low everywhere else.

In Fig. 6 we present, for an average SNR value of $\bar{\gamma} = 10$ dB how the end-to-end backlog increases as a function of the arrival rate $\rho(s)$ for different network sizes. Here, the maximum achievable rate that results in a finite backlog decreases as the number of nodes is increased.

Suppose that buffer sizes are set to satisfy the end-to-end backlog. Then for a fixed buffer size $b_{\max}$, we can use the probability $Pr(B_{\mathrm{net}}(t) > b_{\max})$ as an estimate of the probability of dropped traffic, which we refer to as *loss probability*. In Fig. 7, we depict the loss probability as a function of the average channel SNR for $b_{\max} = 400$ kb, traffic with a rate of $\rho(s) = 20$ and 30 kbps, and for $N = 1, 10,$ and 20 nodes. The figure shows the minimum SNR needed to support a given loss probability is very sensitive to the number of network nodes.

Now we evaluate the violation probability for given end-to-end delay bounds, for a single node ($N = 1$) and a multi-hop

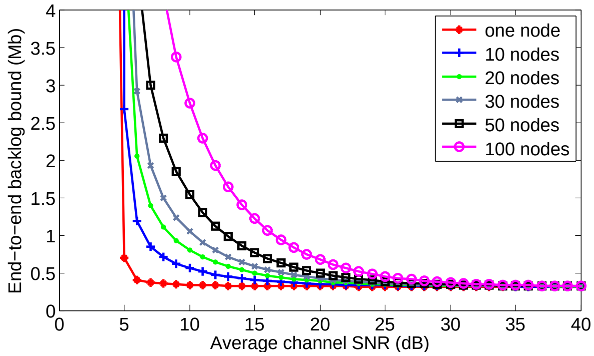

Fig. 5. End-to-end backlog bound ($b^{\varepsilon}_{\text{net}}$) vs. average channel SNR ($\bar{\gamma}$) for multi-hop Rayleigh fading channels with $\varepsilon = 10^{-4}$, $(\sigma(s), \rho(s))$ bounded traffic with $\sigma(s) = 50$ kb and $\rho(s) = 30$ kbps, and $W = 20$ kHz.

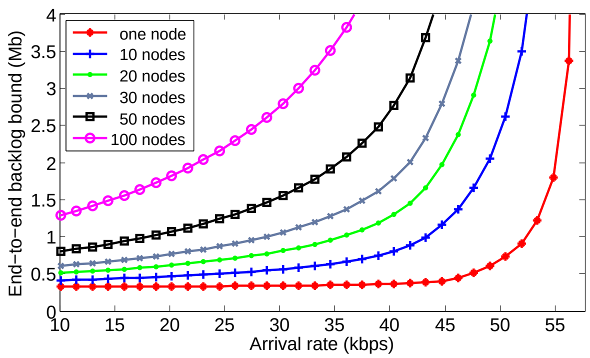

Fig. 6. End-to-end backlog bound (b^{ε}_{net}) vs. arrival rate ($\rho(s)$) for multi-hop Rayleigh fading channels with $\varepsilon = 10^{-4}$, $(\sigma(s), \rho(s))$ bounded traffic with $\sigma(s) = 50$ kb and $\bar{\gamma} = 10$ dB, and $W = 20$ kHz.

network ($N = 10$) for different SNR values. As before, the through traffic is deterministic, with $\sigma(s) = 50$ kb for the burst and $\rho(s) = 20$ kbps for all s. For this example, we also include simulation results. The simulated through traffic consists of a Markov modulated On-Off flow with $\alpha = 0.7$, $\beta = 0.4$, and peak rate $r = 20\frac{\alpha+\beta}{\alpha}$ kbps, which is subsequently shaped by a token bucket with bucket size 50 kb and rate 20 kbps. This creates a bursty traffic flow that saturates the rate of the token bucket. We use the simulations to evaluate the accuracy of our bounds for violation probabilities ranging from 10^0 to 10^{-8}. Fig. 8 illustrates that at sufficiently high SNR values, low delays are achieved even when traffic traverses 10 links. When the SNR is decreased, we observe how the delay performance deteriorates in the multi-hop scenario. The graphs illustrate the dependence of the exponential decay rate of the tail of the delay distribution on the average SNR, i.e., the decay rate increases with $\bar{\gamma}$. A comparison of analytical and simulation results shows that the computed upper bounds provided by our analysis are reasonably close to the simulated system, and reflect the same decay. The results also show that the computed bound are closer to simulation results when the number of nodes is small.

B. Performance Bounds With Cross Traffic

As a final example, we consider Rayleigh fading channels with cross traffic and study the impact of cross traffic char-

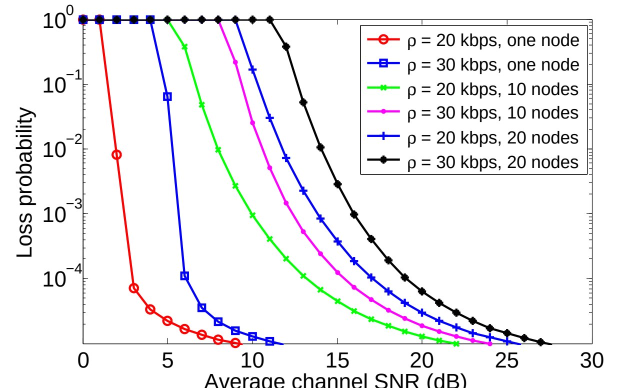

Fig. 7. Loss probability ($\varepsilon(b)$) vs. average channel SNR ($\bar{\gamma}$) for multi-hop Rayleigh fading channels for $N = 1, 10$ and 20 with buffer size 400 kb, $(\sigma(s), \rho(s))$ bounded traffic with $\sigma(s) = 50$ kb and $\rho(s) = 20$ or 30 kbps, and $W = 20$ kHz.

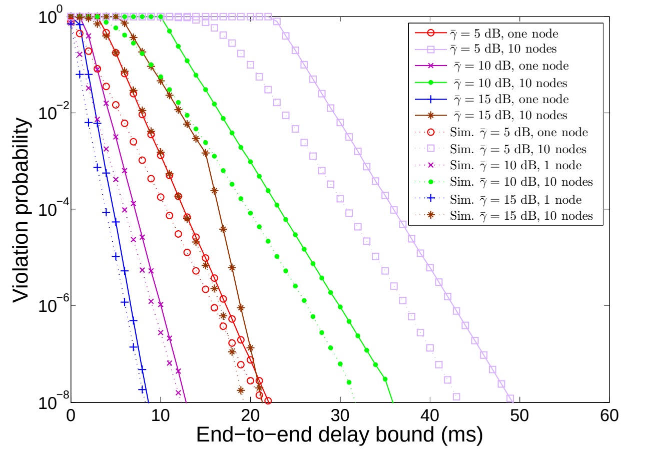

Fig. 8. Delay bound violation probability ($\varepsilon(w)$), computed and simulated, vs. end-to-end delays ($w^{\varepsilon}_{\text{net}}$) for multihop Rayleigh fading channels for $N = 1, 10$, $\bar{\gamma} = 5, 10, 15$ dB, $(\sigma(s), \rho(s))$ bounded traffic with $\sigma(s) = 50$ kb and $\rho(s) = 20$ kbps, and $W = 20$ kHz.

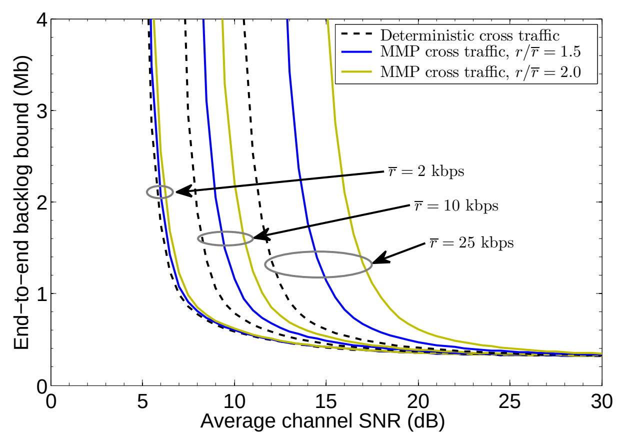

Fig. 9. End-to-end backlog bound ($b^{\varepsilon}_{o,\text{net}}$) vs. average channel SNR ($\bar{\gamma}$) for multi-hop Rayleigh fading channels with Markov modulated On-Off cross traffic. Parameters are $\varepsilon = 10^{-4}$, $N = 10$, $W = 20$ kHz, $(\sigma(s), \rho(s))$ bounded through traffic with $\sigma_o(s) = 50$ kb, $\rho_o(s) = 30$ kbps, cross traffic has average rates $\bar{r} = 2, 10, 25$ kbps and peak-to-average ratios $r/\bar{r} = 1$ (deterministic), 1.5, 2.

acteristics on the channel quality experienced by the through traffic. We fix the number of nodes to $N = 10$. The parameters of the Raleigh channel are as used earlier. The through traffic is again deterministic with the default parameters given at the beginning of the section. The cross traffic is based on a Markov modulated On-Off arrival process as characterized in Subsec. V-A. Its average rate, denoted by $\overline{r}$, is obtained as $\overline{r} = \frac{\alpha}{\alpha+\beta} r$. The parameters α and β are selected so that the average cycle time of the Markov chain as well as the peak-to-average traffic rate have given values. Specifically, we set the cycle time $\frac{1}{\alpha} + \frac{1}{\beta} = 10$ ms, and the peak-to-average ratio to $r/\overline{r} = 1.5$ and 2. The different peak-to-average ratios indicate the burstiness of the cross traffic. We also consider deterministic cross traffic as our baseline for comparison, where $r = \overline{r}$, that is, cross traffic is a constant bit rate traffic source.

In Fig. 9 we show end-to-end backlog bounds $b^{\varepsilon}_{o,\mathrm{net}}$ ($\varepsilon = 10^{-4}$) for the through traffic as a function of the average channel SNR $\bar{\gamma}$. The graphs are grouped according to the average cross traffic rate $\overline{r}$. Obviously, the service offered to the through traffic by the fading channel is reduced when $\overline{r}$ is increased. We observe that for smaller values of $\overline{r}$, varying the burstiness of traffic has a less pronounced effect than for large values of $\overline{r}$. As expected, when the cross traffic has a larger share of the total traffic, its traffic characteristics have a bigger influence on the perceived channel quality of the through traffic. However, the impact of varying the burstiness is remarkable. For $\overline{r} = 10$ and 25, the blow up of the backlog occurs much earlier when the peak-to-average ratio of traffic is increased. Nevertheless, according to Subsec. IV-E, all curves with the same value of $\overline{r}$ have identical asymptotic behavior. We note that with larger peak-to-average ratio values ($r/\overline{r} > 2$) for $\overline{r} = 10$, the backlog bounds will surpass that of $\overline{r} = 25$ with no or moderate burstiness. This provides evidence of the major role of burstiness of interfering flows on the performance of wireless communication channels.

VII. Conclusion

We have developed an analysis of networks with multi-hop fading channels that can incorporate fading channel distributions, without the need for secondary models, such as FSMC. Since such models generally leave open the accuracy of model parameters, they may raise concerns over the fidelity of computed performance metrics with respect to the actual channel. In this paper, we took a fresh point of view, where the descriptions of the arrivals and the fading channels reside in different domains, referred to as bit domain and SNR domain. We found that by mapping arrival processes to the SNR domain, an end-to-end analysis with fading channels becomes tractable. An important discovery was that arrivals and service in the SNR domain obey the laws of a $(\min, \times)$ dioid algebra. The analytical framework developed in this paper appears suitable to study a broad class of fading channels and their impact on the network-layer performance in wireless networks. Even though we made numerous assumptions for the fading channels, our $(\min, \times)$ network calculus may be applicable to networks where these assumptions are relaxed. Generalizing our framework and obtaining a more profound understanding of the dioid algebra and computational methods in the SNR domain is the subject of future research.

Appendix

Lemma 7. *Let $N \geq 1$. For all x with $|x| < 1$,*

$$\sum_{u=0}^{\infty} \binom{N-1+u}{u} x^u = \frac{1}{(1-x)^N} \,. \tag{42}$$

Proof: For $N = 1$, the sum reduces to the geometric series. For $N > 1$, we expand the right hand side as

$$\frac{1}{(1-x)^N} = \left(\sum_{u=1}^{\infty} x^u \right)^N = \sum_{u=0}^{\infty} \Bigg(\sum_{u_1 + \ldots u_N = u} 1 \Bigg) x^u \,.$$

The last sum in parentheses counts the number of N-tuples of nonnegative integers that add up to u. Since its value equals the binomial coefficient $\binom{N-1+u}{u}$, the claim is proved. ■

Lemma 8. *For every $w \geq 0$ and all x with $0 \leq x < 1$,*

$$\sum_{u=w}^{\infty} \binom{N-1+u}{u} x^u \leq \frac{x^w (1+w)^{N-1}}{(1-x)^N} \,.$$

Proof: We write the binomial coefficient as a product,

$$\binom{N-1+u}{u} = \prod_{j=1}^{N-1} \frac{j+u}{j} \,.$$

Since $(j+u) \leq (j+u-w)(1+w)$ for all $j \geq 1$ and all $u \geq w$, it follows that

$$\binom{N-1+u}{u} \leq \binom{N-1+u-w}{u-w} (1+w)^{N-1} \,.$$

This yields for the sum

$$\begin{aligned} \sum_{u=w}^{\infty} \binom{N-1+u}{u} x^u &\leq \sum_{u=w}^{\infty} \binom{N-1+u-w}{u-w} x^u (1+w)^{N-1} \\ &= \frac{1}{(1-x)^N} \, x^w (1+w)^{N-1} \,. \end{aligned}$$

In the last step, we have changed variables to $k = u - w$ and used again Eq. (42). ■

References

[1] H. Al-Zubaidy, J. Liebeherr, and A. Burchard. A (min, x) network calculus for multi-hop fading channels. In *Proc. IEEE Infocom*, 2013.

[2] G. Amarasuriya, C. Tellambura, and M. Ardakani. Asymptotically-exact performance bounds of AF multi-hop relaying over Nakagami fading. *IEEE Trans. on Communications*, 59(4):962–967, 2011.

[3] F. Baccelli, G. Cohen, G. J. Olsder, J. Quadrat. *Synchronization and linearity: an algebra for discrete event systems*. Wiley, 1992.

[4] N. Bisnik and A. A. Abouzeid. Queuing network models for delay analysis of multihop wireless ad hoc networks. *Ad Hoc Networks*, 7(1):79–97, 2009.

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

[6] C.-S. Chang. *Performance guarantees in communication networks*. Springer Verlag, 2000.

[7] C.-S. Chang and R. L. Cruz. A time varying filtering theory for constrained traffic regulation and dynamic service guarantees. In *Proc. IEEE Infocom*, pages 63–70, 1999.

[8] J. Choe and N. B. Shroff. A central-limit-theorem-based approach for analyzing queue behavior in high-speed networks. *IEEE/ACM Trans. on Networking*, 6(5):659–671, 1998.

[9] F. Ciucu. Non-asymptotic capacity and delay analysis of mobile wireless networks. In *Proc. ACM Sigmetrics*, pages 359–360, June 2011.

[10] F. Ciucu, P. Hui, and O. Hohlfeld. Non-asymptotic throughput and delay distributions in multi-hop wireless networks. In *Proc. 48th Annual Allerton Conf. on Communication, Control, and Computing*, pages 662–669, September 2010.

[11] B. Davies. *Integral transforms and their applications*. Springer-Verlag, NY, 1978.

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

[13] M. Fidler. A network calculus approach to probabilistic quality of service analysis of fading channels. In *Proc. IEEE Globecom*, pages 1–6, Nov. 2006.

[14] P. Giacomazzi and G. Saddemi. Bounded-variance network calculus: Computation of tight approximations of end-to-end delay. In *Proc. IEEE ICC*, pages 170–175, May 2008.

[15] M. O. Hasna and M. S. Alouini. Outage probability of multihop transmission over Nakagami fading channels. *IEEE Communications Letters*, 7(5):216–218, 2003.

[16] M. Hassan, M. M. Krunz, and I. Matta. Markov-based channel characterization for tractable performance analysis in wireless packet networks. *IEEE Trans. on Wireless Communications*, 3(3):821–831, 2004.

[17] Y. Jiang and P. J. Emstad. Analysis of stochastic service guarantees in communication networks: a server model. In *Proc. IWQoS, Springer Lecture Notes in Computer Science 3552*, pages 233–245, June 2005.

[18] Y. Jiang and Y. Liu. *Stochastic network calculus*. Springer, 2008.

[19] X. Jin, A.W. Eckford and T.E. Fuja. LDPC codes for non-coherent block fading channels with correlation: analysis and design. *IEEE Trans. on Communications*, 56(1), 70–80, Jan. 2008.

[20] F. Kelly. Notes on effective bandwidths. In *Stochastic Networks: Theory and Applications. (Editors: F.P. Kelly, S. Zachary and I.B. Ziedins) Royal Statistical Society Lecture Notes Series, 4*, pages 141–168. Oxford University Press, 1996.

[21] L. Le and E. Hossain. Tandem queue models with applications to QoS routing in multihop wireless networks. *IEEE Trans. on Mobile Computing*, 7(8):1025–1040, 2008.

[22] L. Le, A. Nguyen, and E. Hossain. A tandem queue model for performance analysis in multihop wireless networks. In *Proc. IEEE WCNC*, pages 2981–2985, March 2007.

[23] J. Le Boudec and P. Thiran. *Network calculus: a theory of deterministic queuing systems for the Internet*. Lecture Notes in Computer Science 2050. Springer, 2001.

[24] C. Li, A. Burchard, and J. Liebeherr. A network calculus with effective bandwidth. *IEEE/ACM Trans. on Networking*, 15(6):1442–1453, 2007.

[25] C. Li, H. Che, and S. Li. A wireless channel capacity model for quality of service. *IEEE Trans. on Wireless Communications*, 6(1):356–366, 2007.

[26] R. Lübben, M. Fidler, and J. Liebeherr. Stochastic bandwidth estimation in networks with random service. *IEEE Trans. on Networking*, DOI: 10.1109/TNET.2013.2261914, To appear.

[27] K. Mahmood, A. Rizk, and Y. Jiang. On the flow-level delay of a spatial multiplexing MIMO wireless channel. In *Proc. IEEE ICC*, pages 1–6, June 2011.

[28] K. Mahmood, M. Vehkaperä, and Y. Jiang. Cross-layer modeling of randomly spread CDMA using stochastic network calculus. *http://arxiv.org/abs/1105.0215*, 2011.

[29] R. McEliece and W.E. Stark. Channels with block interference. *IEEE Trans. on Information Theory*, 30(1), 44–53, 1984.

[30] J. Proakis and M. Salehi. *Digital Communications (5th Ed.)*. McGraw Hill, 2007.

[31] P. Sadeghi, R. Kennedy, P. Rapajic, and R. Shams. Finite-state Markov modeling of fading channels – a survey of principles and applications. *IEEE Signal Processing Magazine*, 25(5):57–80, 2008.

[32] A. Shwartz and A. Weiss. *Large deviations for performance analysis*. Stochastic Modeling Series. Chapman & Hall, 1995.

[33] T. A. Tsiftsis. Performance of wireless multihop communications systems with cooperative diversity over fading channels. *Wiley Int. J. Communication Systems*, 21(5):559–565, 2008.

[34] G. Verticale. A closed-form expression for queuing delay in Rayleigh fading channels using stochastic network calculus. In *Proc. ACM Q2SWinet '09*, pages 8–12, October 2009.

[35] G. Verticale and P. Giacomazzi. An analytical expression for service curves of fading channels. In *Proc. IEEE Globecom*, pages 635–640, November 2009.

[36] H. Wang and N. Moayeri. Finite-state Markov channel - a useful model for radio communication channels. *IEEE Trans. on Vehicular Technology*, 44(1):163–171, 1995.

[37] Q. Wang, D. Wu, and P. Fan. Effective capacity of a correlated Nakagami-m fading channel. *Wireless Communications and Mobile Computing*, 12(14):1225–1238, 2012.

[38] Q. Wang, D. Wu, and P. Fan. Effective capacity of a correlated Rayleigh fading channel. *Wireless Communications and Mobile Computing*, 11(11):1485–1494, 2011.

[39] D. Wu and R. Negi. Effective capacity: a wireless link model for support of quality of service. *IEEE Trans. on Wireless Communications*, 2(4):630–643, 2003.

[40] D. Wu and R. Negi. Effective capacity-based quality of service measures for wireless networks. *J. on Mobile Networking and Applications*, 11(1):91–99, February 2006.

[41] K. Zheng, L. Lei, Y. Wang, Y. Lin, and W. Wang. Quality-of-service performance bounds in wireless multi-hop relaying networks. *IET Communications*, 5(1):71–78, Jan. 2011.

[42] K. Zheng, F. Liu, L. Lei, C. Lin and Y. Jiang. Stochastic performance analysis of a wireless finite–state Markov channel. *IEEE Trans. on Wireless Communications*, 12(2): 782–793, 2012.