On the Recursive Nature of End-to-End Delay Bound for Heterogeneous Wireless Networks

Neda Petreska*, Hussein Al-Zubaidy§, Rudi Knorr*, James Gross§
*Fraunhofer Institute for Embedded Systems and Communication Technologies ESK
{neda.petreska, rudi.knorr}@esk.fraunhofer.de
§School of Electrical Engineering, KTH Royal Institute of Technology
hzubaidy@kth.se, james.gross@ee.kth.se

***Abstract*—Multi-hop wireless networks are increasingly becoming more relevant to current and emerging wireless network deployment. The need for understanding the performance of such networks in order to be able to provide quantifiable end-to-end quality of service is apparent. Until recently, only asymptotic results that describe the scaling of the delay in the size of the network under numerous conformity conditions were available. Recently, a new methodology for wireless networks performance analysis based on stochastic network calculus was presented [1]. This methodology enables the computation of end-to-end probabilistic delay bound of multi-hop wireless networks in terms of the underlying fading channel parameters. However, the approach assumes identically distributed channel gain which applies to a very specific class of networks. In this work, we seek to develop an end-to-end probabilistic delay bound for multi-hop wireless networks with non-identically distributed channel gains. We show that the delay bound for such networks can be computed recursively. We validate the resulting bound by means of simulation and discuss various numerical examples.**

I. Introduction

Performance analysis of multi-hop wireless networks is a challenging task. This is especially true for heterogeneous wireless networks where the possible use of different radio access technologies at different nodes demands various fading and capacity models for the underlying channels.

Recent results [1] provided a methodology based on stochastic network calculus to obtain closed-form expressions describing the probabilistic end-to-end performance bounds for a multi-hop network of independent and identically distributed (i.i.d.) fading channels. We emphasize that the i.i.d. assumption is necessary to obtain the closed-form expression. However, such model is not suitable for the analysis of heterogeneous wireless networks. In fact, such expression cannot be obtained even for homogeneous multi-hop wireless networks with non-identically distributed channel gains.

In this work, we derive an end-to-end probabilistic delay bound for a path consisting of n fading channels with independent yet non-identically distributed channel gains. The relaxation of the identically distributed channel gains in our model makes it well suited to analyze the performance of paths within heterogeneous wireless networks. A direct application of the theory presented in [1] to such scenario requires the determination of $n-1$ computationally intensive convolution operations which may not be feasible for arbitrarily large n. To simplify such complex computations we provide a recursion over the ordered set of links $\mathbb{L}$, that enables the computation of a delay bound for $\mathbb{L}$ from that of $\mathbb{L} \setminus \{n\}$, where n is the last link of the considered path. We prove that the recursive formula we provide does indeed describe the desired end-to-end performance bounds which is furthermore supported by validation done by means of simulation. The recursive structure of the derived expression exposes interesting characteristics of multi-hop wireless networks and quantifies the effect of adding an extra link to an existing path. We believe that this insight might have applications in network planning, admission control, routing and resource allocation for delay sensitive traffic over multi-hop wireless networks.

The rest of the paper is organized as follows. After presenting related work, in Sec. III we discuss the system model and briefly introduce stochastic network calculus. In Sec. IV we derive our main results, while we validate these results in Sec. V. We also point out some insights obtained from numerical results. Finally, we conclude the paper in Sec. VI.

II. Related Work

Delay analysis of wireless multi-hop networks has usually been addressed by using classical queuing theory, e.g., [2]–[4]. Xie and Haenggi [3] investigate the end-to-end delay of a wireless multi-hop line network under both m-phase spatial TDMA and slotted ALOHA. They use a Gl/Geo/1 queuing model. While the authors consider various models for the arrival process, the channel is assumed to offer a Bernoulli service. Hence, the validity and usability of the model is rather limited. Gupta and Shroff [4] use classical queuing theory to derive a lower bound on the average queuing delay of a multi-hop wireless network under pre-specified routing and traffic information. Bisnik and Abouzeid [2] use diffusion approximation to obtain a closed-form expression for the average end-to-end delay in a queuing network model of a random access multi-hop wireless ad hoc network. Using the average service time of the nodes, the authors provide an expression for the achievable throughput. Srinivasa and Haenggi [5] study the throughput-delay-reliability trade-off in multi-hop ad hoc networks using slotted-ALOHA channel access scheme in both noise-limited and interference-limited regimes.

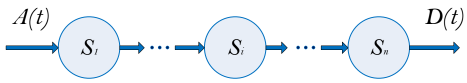

Fig. 1. A model for multi-hop network.

The queuing network based solutions, including the ones mentioned above, rely heavily on conformity assumptions, e.g., regarding traffic and service distributions, which limits their applicability in wireless networks with fading channels. In most cases, the queuing theory approach results in average guarantees on the delay performance. An alternative approach to performance analysis of multi-hop wireless networks is the stochastic network calculus [6]. Extensive work on end-to-end delay analysis using stochastic network calculus can be found in the literature. More recently, it is being applied to the performance analysis of wireless networks, e.g., [7]–[11]. Zheng et al. [7] conducted delay and backlog analysis of wireless Rayleigh-fading channels, using a finite-state Markov channel model without considering multi-hop paths. Ciucu et al. [8], [9] provide non-asymptotic throughput and delay bounds for multi-hop wireless networks. The node capacity is assumed to be limited by the interference from other transmitters in the network, i.e., slotted-ALOHA. Ciucu et al. [10] extends this analysis and shows how to fit MAC protocols into the proposed methodology, highlighting when multi-hop routing is more advantageous than single-hop routing. However, none of [8]–[10] incorporates the underlying fading channel characteristics. Fidler [11] investigates probabilistic performance guarantees of wireless fading channels using moment generating function-based stochastic network calculus. The underlying physical channel model is represented by the two state Gilbert-Elliott model. Although the proposed solution represents a useful model to derive backlog and delay bounds for wireless links, due to the complexity of the service model, it is not feasible to be applied to multi-hop networks.

III. PRELIMINARIES

In this section we present the basic system model as well as the problem statement. A brief overview of stochastic network calculus, as a foundation for our derivations, is also provided.

A. System Model and Problem Statement

We consider a single traffic flow, or an aggregate of flows, traversing a multi-hop path consisting of n buffered wireless links as shown in Fig.1. The path is represented by an ordered set $\mathbb{L} = \{1, \ldots, n\}$. Time is slotted with periods T. We assume a fluid flow traffic model. Per time slot i a random amount of bits a_i arrive to the first link and are stored in its queue until transmission. Furthermore, per time slot i each link j can transport a randomly varying amount of $c_{i,j}$ bits. Data arriving at the last node departs from the system (i.e. it is passed to the application etc.) without any additional delay.

The random service process $c_{i,j}$ of each link is a result of the randomly varying signal-to-noise ratio $\gamma_{i,j}$ (SNR) at the receiver of the corresponding link. Assuming link j to be operated with a transmit power of P_{tx}, the SNR is obtained as $\gamma_{i,j} = \frac{P_{\mathrm{tx}} \cdot h_{i,j}^2}{\sigma^2}$, where $h_{i,j}^2$ denotes the instantaneous channel gain and σ^2 denotes the noise power. The channel gain $h_{i,j}^2$ is a random variable and is assumed in the following to have an arbitrarily different distribution per link. A common assumption for the channel gain is Rayleigh fading, resulting in an exponentially distributed channel gain, and therefore also an exponentially distributed SNR $\gamma_{i,j}$ with average $\bar{\gamma}_j$. A block-fading model [12] is assumed in the following, i.e. the instantaneous SNR remains constant within a given time slot, while it varies independently from one time slot to the next. In addition, we assume the SNR of different links to be statistically independent. Given this random SNR $\gamma_{i,j}$, the link service capacity $c_{i,j}$ follows from Shannon's formula as $c_{i,j} = N \cdot \log_2(1 + \gamma_{i,j})$, where N denotes the number of symbols that can be transmitted during one time slot.

We define $W(t)$ as the random delay, which bits entering the system at slot t will experience while being conveyed to the destination. In the following we are interested in a probabilistic bound on this end-to-end delay and will derive it based on stochastic network calculus.

B. Stochastic Network Calculus

The delay analysis of a multi-hop path with queuing effects has traditionally been a hard problem. As we are mainly interested in a stochastic behavior of the links (due to our interest in wireless networks), we restrict the discussion in the following to stochastic network calculus [6], [13]. In this theory, a queuing network with stochastic arrival and departure process is considered, where the bivariate functions $A(\tau, t)$, $D(\tau, t)$ and $S(\tau, t)$ (for any $0 \le \tau \le t$) denote the *cumulative* arrival to the system, departure from the system as well as service of the system, respectively. More precisely, let us assume for now a single wireless link. Denote by d_i the stochastic departure of bits from the system during time slot i (i.e. the bits that are leaving the system), while a_i and c_i denote the (stochastic) arrival to and the (stochastic) service offered by the system, respectively, during time slot i. The cumulative functions follow as $A(\tau, t) = \sum_{i=\tau}^{t-1} a_i$ for the arrival process, $D(\tau, t) = \sum_{i=\tau}^{t-1} d_i$ for the departure process and $S(\tau, t) = \sum_{i=\tau}^{t-1} c_i$ for the service process.

Stationary performance bounds can be obtained only when the analyzed system satisfies a stability condition. A queuing system is stable when the average arrival rate is smaller than the average service rate, that is:

$$\lim_{t \to \infty} \frac{A(0,t)}{t} < \lim_{t \to \infty} \frac{S(0,t)}{t}.$$

In our analysis, $S(\tau, t)$ is a *dynamic server* satisfying:

$$D(0,t) \ge A \otimes S(0,t),$$

where $\otimes$ is the $(\min, +)$ convolution operator[1], defined as

[1]Network calculus is a system-theoretic interpretation of the dynamics of a queuing system based on $(\min, +)$ algebra. This unusual algebra is a significant obstacle to overcome when trying to understand the theory.

$x \otimes y(\tau,t) = \inf_{t\geq u\geq\tau}\{x(\tau,u)+y(u,t)\}$. Based on this, we also define the $(\min,+)$ deconvolution as $x \oslash y(\tau,t) = \sup_{\tau\geq u\geq 0}\{x(u,t)-y(u,\tau)\}$. Under these assumptions we are interested in the stochastic delay $W(t)$ of the system at time t, which directly results from the definition of the cumulative arrival and departure:

$$W(t) \leq \inf\{u \geq 0 : A \oslash S(t+u,t) \leq 0\}\,. \tag{1}$$

More precisely, as $A(\tau,t)$ and $D(\tau,t)$ are stochastic bivariate functions, we are interested in a probabilistic bound on $W(t)$ in the form of $\Pr[W(t) > w^{\varepsilon}] \leq \varepsilon$, which is also known as the *violation probability* for a target delay w^{ε}. Such a bound can be found based on the moment generating function (MGF) of the cumulative arrival and service processes for any θ [13]:

$$\mathsf{M}_{A(\tau,t)}(\theta) = \mathrm{E}\left[e^{\theta A(\tau,t)}\right], \mathsf{M}_{S(\tau,t)}(\theta) = \mathrm{E}\left[e^{\theta S(\tau,t)}\right].$$

A bound on the delay as given by Eq. (1) follows from a bound on the deconvolution of the moment generating functions [13]. However, determining the MGF of the cumulative service process of wireless systems has been found to be a notoriously difficult problem, as also witnessed in the context of the effective service capacity of wireless systems [14].

Recently, a more promising approach for wireless networks has been proposed in [1], where the queuing behavior is analyzed directly in the "domain" of channel variations instead of the bit domain. This can be interpreted as the *SNR domain* (thinking of bits as "SNR demands" that reside in the system until these demands can be met by the channel), in contrast to the *bit domain* addressed by the MGF-based analysis. To start with, the cumulative arrival, service and departure processes in the bit domain, i.e., A, D and S, are related to their SNR domain counterparts (represented in the following by calligraphic upper case letters $\mathcal{A}, \mathcal{D}$ and $\mathcal{S}$ respectively), through the exponential function. Thus, we have $\mathcal{A}(\tau,t) \triangleq e^{A(\tau,t)}, \mathcal{D}(\tau,t) \triangleq e^{D(\tau,t)}$ and $\mathcal{S}(\tau,t) \triangleq e^{S(\tau,t)}$. Due to the exponential function, these cumulative functions become products of the increments in the bit domain, i.e., for the cumulative service process in the SNR domain we have:

$$\mathcal{S}(\tau,t) = \prod_{i=\tau}^{t-1} e^{c_i} = \prod_{i=\tau}^{t-1}(1+\gamma_i)^{\mathcal{N}} = \prod_{i=\tau}^{t-1} g(\gamma_i)\,,$$

where $\mathcal{N} = N/\ln 2$. Furthermore, the delay at time t is obtained in analogy to Eq. (1)[2]:

$$\mathcal{W}(t) = W(t) \leq \inf\{u \geq 0 : \mathcal{A} \oslash \mathcal{S}(t+u,t) \leq 1\}\,.$$

As with the MGF-based analysis approach, a bound ε for the delay violation probability $\Pr[W(t) > w^{\varepsilon}]$ can be derived based on a transform of the cumulative arrival and service processes in the SNR domain. In [1] it was shown that such a violation probability bound for a given w^{ε} must satisfy:

$$\inf_{s>0}\{\mathcal{K}(s,t+w^{\varepsilon},t)\} \leq \varepsilon\,. \tag{2}$$

[2]In the SNR domain the system-theoretic interpretation of the queuing dynamics is based now on $(\min,\times)$ algebra due to the exponential function in the definition of the cumulative arrival, service and departure processes.

We refer to the function $\mathcal{K}(s,\tau,t)$ as *kernel* defined as:

$$\mathcal{K}(s,\tau,t) = \sum_{i=0}^{\min(\tau,t)} \mathcal{M}_{\mathcal{A}}(1+s,i,t)\cdot\mathcal{M}_{\mathcal{S}}(1-s,i,\tau). \tag{3}$$

The function $\mathcal{M}_X(s)$ is the Mellin transform [15] of a random process, defined as:

$$\mathcal{M}_X(s,\tau,t) = \mathcal{M}_{X(\tau,t)}(s) = \mathrm{E}\left[X^{s-1}(\tau,t)\right], s \in \mathbb{R}\,.$$

In the following, we will assume $\mathcal{A}(\tau,t)$ and $\mathcal{S}(\tau,t)$ to have stationary increments. We denote them by α for the arrivals (in SNR domain) and $g(\gamma)$ for the service. Hence, the Mellin transform becomes independent of the time instance, which we account for by denoting $\mathcal{M}_X(s,\tau,t) = \mathcal{M}_X(s,t-\tau)$. In addition, as we only consider stable queuing systems in a steady state, the kernel becomes independent of the time instance t and we denote $\mathcal{K}(s,t+w,t) \overset{t\to\infty}{=} \mathcal{K}(s,-w)$.

The strength of the Mellin-transform-based approach becomes apparent when considering block-fading channels. The Mellin transform for the cumulative service process in SNR domain is given by:

$$\mathcal{M}_{\mathcal{S}}(s,\tau,t) = \prod_{i=\tau}^{t-1}\mathcal{M}_{g(\gamma)}(s) = \mathcal{M}_{g(\gamma)}^{t-\tau}(s) = \mathcal{M}_{\mathcal{S}}(s,t-\tau)\,,$$

where $\mathcal{M}_{g(\gamma)}(s)$ is the Mellin transform of the (stationary) service increment $g(\gamma)$ in the SNR domain. The function $g(\cdot)$ represents here the modification of the SNR due to the Shannon formula. However, it can also model more complex system characteristics, most importantly scheduling effects.

Assuming Rayleigh fading, i.e., an exponentially distributed SNR at the receiver, the Mellin transform of the service will result into:

$$\mathcal{M}_{g(\gamma)}(s) = \left[e^{1/\bar{\gamma}}\cdot\bar{\gamma}^{\mathcal{N}(s-1)}\Gamma\left(\mathcal{N}(s-1)+1, 1/\bar{\gamma}\right)\right],$$

where $\Gamma(x,y) = \int_y^{\infty} t^{x-1}e^{-t}dt$ is the incomplete gamma function.

Assuming the cumulative arrival process in SNR domain to have independent increments and denoting its corresponding Mellin transform by $\mathcal{M}_{\mathcal{A}}(s,t-\tau) = \prod_{i=\tau}^{t-1}\mathcal{M}_{\alpha}(s)$, the steady-state kernel for a Rayleigh-fading wireless channel results into:

$$\mathcal{K}(s,-w) = \frac{\mathcal{M}_{g(\gamma)}^{w}(1-s)}{1-\mathcal{M}_{\alpha}(1+s)\cdot\mathcal{M}_{g(\gamma)}(1-s)} = \frac{\left(e^{1/\bar{\gamma}}\cdot\bar{\gamma}^{-s\mathcal{N}}\cdot\Gamma(1-s\mathcal{N},\frac{1}{\bar{\gamma}})\right)^{w}}{1-\mathcal{M}_{\alpha}(1+s)\cdot e^{1/\bar{\gamma}}\cdot\bar{\gamma}^{-s\mathcal{N}}\cdot\Gamma(1-s\mathcal{N},\frac{1}{\bar{\gamma}})} \tag{4}$$

for any $s > 0$ under the stability condition:

$$\mathcal{M}_{\alpha}(1+s)\cdot\mathcal{M}_{g(\gamma)}(1-s) < 1\,. \tag{5}$$

IV. Probabilistic Delay Bound

In this section we present our main result. A major strength of network calculus is the ability to extend one-hop results to multi-hop paths with reasonable effort. In the context of wireless systems, end-to-end probabilistic bounds have only been obtained for concatenated i.i.d. service processes (i.e. for multi-hop wireless links which are all having the same fading process with the same average SNR). In the following we relax this condition and generalize the so far known results to arbitrarily distributed random services per link.

In order to determine an end-to-end delay bound of a path $\mathbb{L} = \{1, \ldots, n\}$, one has to consider the Mellin transform of the joint probabilistic service $\mathcal{S}^{\mathbb{L}}(\tau, t) = \mathcal{S}_1 \otimes \ldots \otimes \mathcal{S}_n$ in the SNR domain.[3] Once that Mellin transform can be determined, the delay bound follows from Eq. (2) and (3) [1]:

Lemma 1. *For a path $\mathbb{L}$ a probabilistic end-to-end delay bound is given by the minimum w^ε that satisfies*

$$\inf_{s>0} \left\{ \mathcal{K}^{\mathbb{L}}(s, -w^\varepsilon) \right\} \leq \varepsilon \,.$$

In order to obtain the Mellin transform of the end-to-end service, consider the following relationship. 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

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

Hence, the corresponding Mellin transform of the path $\mathbb{L}$ can be bounded by [1]:

$$\mathcal{M}_{\mathcal{S}^{\mathbb{L}}}(s, \tau, t) \leq \sum_{i_1=i_0}^{t} \sum_{i_2=i_1}^{t} \cdots \sum_{i_{n-1}=i_{n-2}}^{t} \mathcal{M}_{\mathcal{S}_1}(i_1 - i_0) \cdot \mathcal{M}_{\mathcal{S}_2}(i_2 - i_1) \ldots \mathcal{M}_{\mathcal{S}_n}(i_n - i_{n-1}) = \sum_{i_1 \ldots i_{n-1}}^{t} \prod_{j=1}^{n} \mathcal{M}_{g(\gamma_j)}^{i_j - i_{j-1}}(s) \,, \tag{6}$$

with $\tau = i_0 \leq i_1 \leq \cdots \leq i_n = t$. Notice that $\mathcal{M}_{g(\gamma_j)}(s)$ denotes the Mellin transform of the (stationary) SNR domain service increment of link j.

Based on this relationship, we present now the main contribution of the paper. Let us define $\mathcal{K}^{\mathbb{L}}$ as the kernel for a path $\mathbb{L}$ containing n links. Let $m \in \{1, 2, \ldots, n-1\}$ and n refer to the m^{th} and the n^{th} link, respectively, in the path $\mathbb{L}$.

Theorem 1. *Given a path $\mathbb{L} \setminus \{n\}$ of links with independent and non-identically distributed service processes, with kernel $\mathcal{K}^{\mathbb{L}\setminus\{n\}}$; then $\mathcal{K}^{\mathbb{L}}$ can be obtained in terms of $\mathcal{K}^{\mathbb{L}\setminus\{n\}}$ as follows*

$$\begin{aligned}
&\mathcal{K}^{\mathbb{L}}(s, -w) = \\
&= \frac{\mathcal{M}_{g(\gamma_n)}(1-s)}{\mathcal{M}_{g(\gamma_n)}(1-s) - \mathcal{M}_{g(\gamma_m)}(1-s)} \mathcal{K}^{\mathbb{L}\setminus\{m\}}(s, -w) \\
&+ \frac{\mathcal{M}_{g(\gamma_m)}(1-s)}{\mathcal{M}_{g(\gamma_m)}(1-s) - \mathcal{M}_{g(\gamma_n)}(1-s)} \mathcal{K}^{\mathbb{L}\setminus\{n\}}(s, -w)
\end{aligned}$$

for any $m \in \{1, 2, \ldots, n-1\}$.

Proof. We start by considering the bound on the Mellin transform of the service curve of path $\mathbb{L}$ as given by Eq. (6) with $i_0 = \tau$ and $i_n = t$. Without loss of generality, let $m = n-1$. So, we obtain:

$$\begin{aligned}
&\mathcal{M}_{\mathcal{S}^{\mathbb{L}}}(s, t-\tau) \leq \sum_{i_1,\ldots,i_{n-1}}^{t} \prod_{j=1}^{n} \mathcal{M}_{g(\gamma_j)}^{i_j - i_{j-1}} \\
&= \sum_{i_1,\ldots,i_{n-2}}^{t} \mathcal{M}_{g(\gamma_1)}^{i_1-\tau} \mathcal{M}_{g(\gamma_2)}^{i_2-i_1} \cdots \frac{\mathcal{M}_{g(\gamma_n)}^{t}}{\mathcal{M}_{g(\gamma_{n-1})}^{i_{n-2}}} \sum_{i_{n-1}=i_{n-2}}^{t} \left(\frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}} \right)^{i_{n-1}} \\
&= \sum_{i_1,\ldots,i_{n-2}}^{t} \mathcal{M}_{g(\gamma_1)}^{i_1-\tau} \cdot \mathcal{M}_{g(\gamma_2)}^{i_2-i_1} \cdots \frac{\mathcal{M}_{g(\gamma_n)}^{t}}{\mathcal{M}_{g(\gamma_{n-1})}^{i_{n-2}}} \\
&\cdot \left(\frac{\left(\frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}}\right)^{i_{n-2}} - \left(\frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}}\right)^{t+1}}{1 - \frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}}} \right) \\
&= \frac{\mathcal{M}_{g(\gamma_n)}}{\mathcal{M}_{g(\gamma_n)} - \mathcal{M}_{g(\gamma_{n-1})}} \sum_{i_1,\ldots,i_{n-2}}^{t} \mathcal{M}_{g(\gamma_1)}^{i_1-\tau} \cdots \frac{\mathcal{M}_{g(\gamma_n)}^{t}}{\mathcal{M}_{g(\gamma_{n-1})}^{i_{n-2}}} \\
&\cdot \left(\frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}} \right)^{i_{n-2}} - \frac{\mathcal{M}_{g(\gamma_n)}}{\mathcal{M}_{g(\gamma_n)} - \mathcal{M}_{g(\gamma_{n-1})}} \\
&\sum_{i_1,\ldots,i_{n-2}}^{t} \mathcal{M}_{g(\gamma_1)}^{i_1-\tau} \cdots \frac{\mathcal{M}_{g(\gamma_n)}{}^{t}}{\mathcal{M}_{g(\gamma_{n-1})}^{i_{n-2}}} \left(\frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}} \right)^{t} \frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)}} \\
&= \frac{\mathcal{M}_{g(\gamma_n)}}{\mathcal{M}_{g(\gamma_n)} - \mathcal{M}_{g(\gamma_{n-1})}} \sum_{i_1,\ldots,i_{n-2}}^{t} \mathcal{M}_{g(\gamma_1)}^{i_1-\tau} \cdots \mathcal{M}_{g(\gamma_{n-2})}^{i_{n-3}-i_{n-2}} \mathcal{M}_{g(\gamma_n)}^{t-i_{n-2}} \\
&- \frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_n)} - \mathcal{M}_{g(\gamma_{n-1})}} \sum_{i_1,\ldots,i_{n-2}}^{t} \mathcal{M}_{g(\gamma_1)}^{i_1-\tau} \cdots \mathcal{M}_{g(\gamma_{n-2})}^{i_{n-3}-i_{n-2}} \mathcal{M}_{g(\gamma_{n-1})}^{t-i_{n-2}} \\
&= \frac{\mathcal{M}_{g(\gamma_n)}}{\mathcal{M}_{g(\gamma_n)} - \mathcal{M}_{g(\gamma_{n-1})}} \mathcal{M}_{\mathcal{S}^{\mathbb{L}\setminus\{n-1\}}}(t-\tau) \\
&+ \frac{\mathcal{M}_{g(\gamma_{n-1})}}{\mathcal{M}_{g(\gamma_{n-1})} - \mathcal{M}_{g(\gamma_n)}} \mathcal{M}_{\mathcal{S}^{\mathbb{L}\setminus\{n\}}}(t-\tau) \,,
\end{aligned}$$

where we have omitted for readability that all Mellin transforms above are functions of s. Thus, we have shown that an upper bound of the Mellin transform of path $\mathbb{L}$ can be obtained recursively from the Mellin transform of the service process of path $\mathbb{L} \setminus \{n-1\}$ and $\mathbb{L} \setminus \{n\}$.

As the kernel is a function of the Mellin transforms of the SNR domain arrival and service process, i.e.,

$$\mathcal{K}^{\mathbb{L}}(s, t+w, t) = \sum_{i=0}^{t} \mathcal{M}_{\mathcal{A}}(1+s, i, t) \, \mathcal{M}_{\mathcal{S}^{\mathbb{L}}}(1-s, i, t+w) \,,$$

it follows directly that the steady state kernel $\mathcal{K}^{\mathbb{L}}(s, -w)$ is a recursive function of the kernels $\mathcal{K}^{\mathbb{L}\setminus\{n-1\}}(s, -w)$ and $\mathcal{K}^{\mathbb{L}\setminus\{n\}}(s, -w)$ as claimed in the theorem. $\square$

[3] In the following $\otimes$ denotes the convolution in $(\min, \times)$ algebra.

A direct consequence of Theorem 1 and Lemma 1 is obviously that the delay bound for path $\mathbb{L}$ can be obtained from recursively computing the kernel according to the theorem. In this recursion, the number of summands increases with the number of hops. For an n-hop path there are 2^{n-1} summands, as each geometric sum results into two summands. Although the computational complexity of the proposed recursive formula grows geometrically in n, its performance still outperforms the alternative where n convolution processes need to be evaluated in order to obtain the end-to-end network performance. The computational complexity can be further reduced for scenarios where the analyzed heterogeneous path is composed of $h \leq n$ segments that are themselves composed of groups of homogeneous links. In this case, we propose a hybrid approach where $\mathcal{M}_{g(\gamma_k)}, k = 1, \ldots h$, is computed using the closed-form formula for homogeneous network presented in [1], then $\mathcal{M}_{\mathcal{S}^{\mathbb{L}}}$ is computed using Theorem 1. Then the complexity is reduced by a factor of $\frac{n-h}{n}$.

Note further, that the stability condition in Eq. (5) needs to hold for all individual links:

$$\max_j \left(\mathcal{M}_{\alpha}\left(1+s\right) \cdot \mathcal{M}_{g(\gamma_j)}\left(1-s\right)\right) < 1.$$

We next discuss the interpretation of Theorem 1 and its implications on the performance analysis of multi-hop wireless networks. Theorem 1 suggests that the delay bound is increased each time a new link is added to the path under investigation. This increase depends on how the added wireless link's SNR distribution compares to the SNR distributions of the links already incorporated in that path, i.e., an added link with worse SNR distribution will contribute significantly to the delay bound and vice versa. Furthermore, Theorem 1 submits that this increase is quantifiable and it provides a recursive formula to compute the probabilistic end-to-end delay bound of the n-hop path in terms of that of the $n-1$-hop path. This characteristic of the bound facilitates the determination of a best route, based on wireless link quality defined by the SNR distributions, by building a path in a stepwise fashion using the recursive formula given in Theorem 1. This may allow the development of efficient wireless routing algorithms considering a target delay with corresponding violation probability.

To show how Theorem 1 can be used in practice, suppose that we have determined the probabilistic delay bound of a multi-hop path consisting of wireless links $1, 2, ..., n-1$ and want to compute the probabilistic delay bound of an n-hop path by adding an additional link, n, at the end of that path. In order to do this we have to compute the sum of the delay kernel of two $n-1$-hop sub-paths. The first of the two terms is the delay bound kernel of a path which does not contain the last link n. This function represents the probability that a delay bound w^{ε} is violated, in the sense of Lemma 1, for a path $\mathbb{L} \setminus \{n\} = \{1, 2, ..., n-1\}$. The second term, is the same as the first term except that the m^{th} link, $m \in \{1, 2, ..., n-1\}$, is replaced with the n^{th} link, i.e., it is the kernel for the path $\mathbb{L} \setminus \{m\}$. An interesting feature of the recursive bound is the fact that we can choose *any* m. Without loss of generality, we choose $m = n-1$ as given in the proof of Theorem 1.

Furthermore, each of the two terms is scaled by a factor that is a function of the n^{th} and the $n-1^{\text{st}}$ channels' fading distributions. These factors can be interpreted as a "price to pay" for replacing link n with link $n-1$ in the $\mathbb{L} \setminus \{n\}$ sub-path. A closer inspection of these two factors reveals the following: Adding a link that is characterized with a more favorable SNR distribution, hence, offering a higher channel capacity, contributes less to the end-to-end delay bound along the path. On the other hand, an added link with a worse SNR distribution contributes more to that delay bound. These values are quantified in terms of the service process Mellin transform of the added link. Suppose that the fading processes of links n and $n-1$ belong to the same distribution family, e.g., Rayleigh, and have $\bar{\gamma}_n < \bar{\gamma}_{n-1}$. In this case $\mathcal{M}_{g(\gamma_n)}(1-s) > \mathcal{M}_{g(\gamma_{n-1})}(1-s)$ and $\frac{\mathcal{M}_{g(\gamma_n)}(1-s)}{\mathcal{M}_{g(\gamma_n)}(1-s) - \mathcal{M}_{g(\gamma_{n-1})}(1-s)} > 1$, so the delay bound of the sub-path excluding the link with the higher SNR will be scaled up and therefore will be emphasized more in the computation of the delay bound for path $\mathbb{L}$. The coefficient $\frac{\mathcal{M}_{g(\gamma_{n-1})}(1-s)}{\mathcal{M}_{g(\gamma_{n-1})}(1-s) - \mathcal{M}_{g(\gamma_n)}(1-s)} < 0$ and the delay bound of the sub-path including the stronger link will be emphasized less in the computation, leading to a smaller reduction in the end-to-end delay bound, which results in turn in a higher overall increase of the delay bound when adding a relatively weak link to the path instead of a stronger one.

V. Validation and Numerical Results

In this section, we validate our analysis by simulations. We also present numerical results for several multi-hop wireless network examples and discuss these results.

A. Analytical Examples

We use Theorem 1 to compute probabilistic delay bounds for n-hop wireless networks, where the underlying channels experience Rayleigh fading. According to the theorem, for a path $\mathbb{L}$ consisting of links 1 and 2 with an average SNR of $\bar{\gamma}_1$ and $\bar{\gamma}_2$ respectively, we get the following delay bound kernel:

$$\mathcal{K}^{\mathbb{L}}(s, -w) = \frac{\mathcal{M}_{g(\gamma_2)}(1-s)}{\mathcal{M}_{g(\gamma_2)}(1-s) - \mathcal{M}_{g(\gamma_1)}(1-s)} \cdot \mathcal{K}^{\{2\}}(s, -w) + \frac{\mathcal{M}_{g(\gamma_1)}(1-s)}{\mathcal{M}_{g(\gamma_1)}(1-s) - \mathcal{M}_{g(\gamma_2)}(1-s)} \cdot \mathcal{K}^{\{1\}}(s, -w), \tag{7}$$

where $\mathcal{K}^{\{i\}}(s, -w), i \in \{1, 2\}$, is the single link kernel defined by Eq. (4). In case of a 3-hop path, we extend Eq. (7) for a path $\mathbb{L}' = \{1, 2, 3\}$ to:

$$\mathcal{K}^{\mathbb{L}'}(s, -w) = \frac{\mathcal{M}_{g(\gamma_3)}(1-s) \cdot \mathcal{K}^{\{1,3\}}(s, -w)}{\mathcal{M}_{g(\gamma_3)}(1-s) - \mathcal{M}_{g(\gamma_2)}(1-s)} + \frac{\mathcal{M}_{g(\gamma_2)}(1-s) \cdot \mathcal{K}^{\{1,2\}}(s, -w)}{\mathcal{M}_{g(\gamma_2)}(1-s) - \mathcal{M}_{g(\gamma_3)}(1-s)},$$

where $\mathcal{K}^{\{1,2\}}(s, -w)$ and $\mathcal{K}^{\{1,3\}}(s, -w)$ are in turn computed by using Eq. (7). In the following numerical results, we have computed the kernel in this recursive manner.

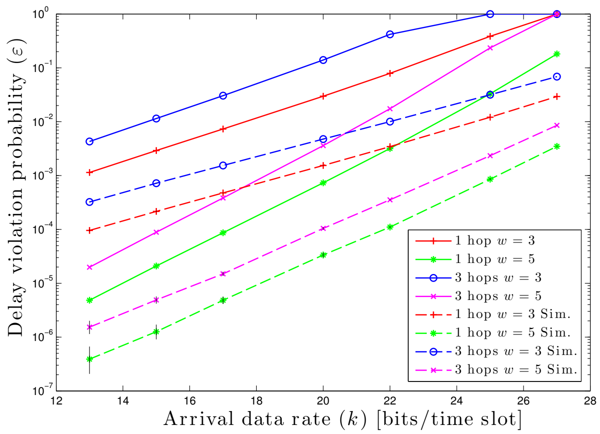

Fig. 2. Analytical delay bound and simulation results for a path consisting of one link with average SNR of 5 dB and a path consisting of 3 links with average SNR of 5, 10 and 7 dB. The target delay is $w = \{3, 5\}$ time slots.

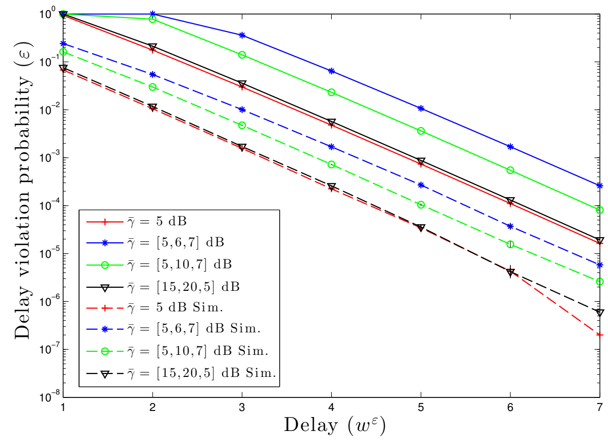

Fig. 3. Three 3-hop paths having the same bottleneck link are compared to one another and to a 1-hop path regarding both the analytical bound and simulation results. The arrival data rate is $k = 20$ bits per time slot.

B. Simulation Setup and Methodology

We simulated wireless paths with various number of hops, where each hop is characterized by a Rayleigh-fading channel with a different average SNR. The instantaneous SNR in each time slot is drawn accordingly from an exponential distribution. The transmission rate of each link is equivalent to the theoretical Shannon capacity as described in Section III-A. The number of symbols per time slot is set to $N = 20$. We use a fluid flow model of the through traffic, entering the system at the first node with a constant data rate of k arriving bits per time slot. During the simulation the end-to-end delay of the departing bits from the last link is sampled and collected to characterize the violation of a delay target, taking into account the correlation among the delay samples. Based on this output, confidence intervals on the violation probability for all simulation runs are obtained with a confidence level of 0.95, but are often not shown, due to their low absolute range.

C. Numerical Results

We first discuss our validation results presented in Figure 2 and 3. Figure 2 compares both simulation and analytical results for the delay violation probability of a 1-hop and a 3-hop path for different arrival rates. The average SNR of the links is $\bar{\gamma} = 5$ dB in the 1-hop case and $\bar{\gamma} = \{5, 10, 7\}$ dB in the 3-hop case. The two cases are compared with respect to an end-to-end delay of 3 and 5 time slots. The figure reveals that the analytical bound is indeed an upper bound to the system performance. The difference between the computed bound and the simulation is around one order of magnitude. This difference is mainly due to the union bound, used to replace the supremum with a sum in the computation of the end-to-end convolution of cascade of links. Figure 2 also reveals that the union bound becomes less tight as k increases. Obviously, for both analytical and simulation results, the violation probability increases as the target delay bound is reduced. It also increases for bigger number of hops.

In Figure 3 we show the resulting violation probability versus the target delay for different constellations of 3-hop paths, each of them having a weakest link with an average SNR of 5 dB. We compare our results to simulations. For different composition of the paths we consider cases with an average SNR of $5, 6, 7$ dB, $5, 10, 7$ dB and $5, 15, 20$ dB. In addition, we provide in the figure the results for a 1-hop link with an average SNR of 5 dB, while the arrival rate is fixed to $k = 20$ bits per time slot. Again the figure validates the analysis in the sense that the derived bound is indeed an upper bound to the simulation results. The results show nicely the expected exponential decay of the tail of the violation probabilities. Moreover, while for the analytical results as well as the simulation results there is hardly any difference in the delay distribution of the 1-hop case versus the 3-hop case with average link SNRs of $5, 15, 20$ dB, the results profoundly differ when comparing the 1-hop link and the 3-hop path with SNRs of $5, 10, 7$ and $5, 6, 7$ dB, with the latter having the highest violation probabilities with respect to the bound as well as the simulations. Figure 3 quantifies the effect of a bottleneck wireless link in a multi-hop path on the network performance and clearly shows that it is affected by other links' qualities as well as the bottleneck link. The effect of the other links on the delay diminishes as the SNR gap to the bottleneck increases.

Figure 4 shows the analytical delay bound for different number of hops n for various delays w^ε. We start with a 1-hop path and then increase the number of hops up to a 7-hop path, each time adding a new link which has either the lowest or the highest average SNR among the links. Hence, we always get a pair of paths that have the same bottleneck link, except for the 1-hop case (see the legend of the figure for an exact specification of the average links' SNR). The figure reveals how the violation probabilities increase every time a weak link is added, in contrast to adding strong links which have only a marginal impact on the violation probability.

Finally, Figure 5 illustrates how different the links in a given path should be, so that the delay bound computed with the recursive formula differs significantly from the simpler one applied for homogeneous links, as provided in [1]. A 3-hop path whose links have equal path loss, is considered. In the

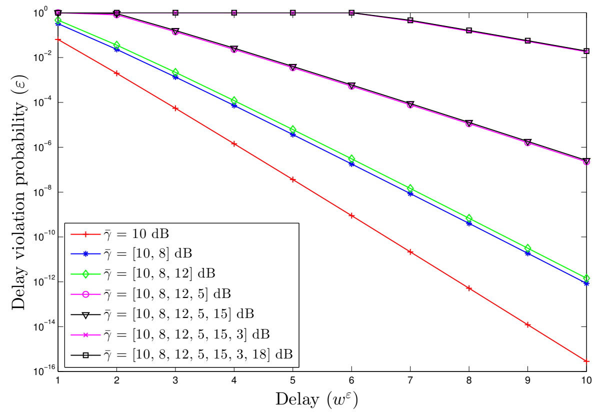

Fig. 4. Analytical delay bound versus the target delay for various paths with up to 7 hops. The arrival rate is fixed at $k = 20$ bits per time slot.

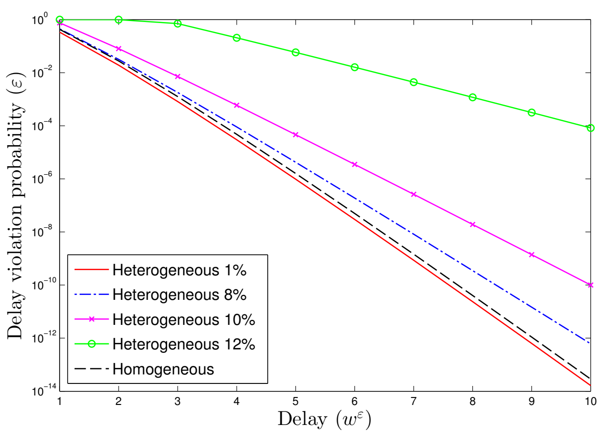

Fig. 5. Comparison between the delay bound function of heterogeneous and homogeneous links for different target delays. The difference between the links is represented in % w.r.t the value used for the homogeneous case. The arrival rate is $k = 22$ bits per time slot.

homogeneous case every node uses a transmit power of 1 mW. Heterogeneous links are created by decreasing and increasing the transmit power for an equal amount on the first and last link, respectively, while maintaining the same total power budget. As a result to the union bound used for obtaining the delay bound function for homogeneous links, we notice the existence of a tighter delay bound in the case of 1% different transmit power. We can generally conclude that in case of more than 10% link divergence, the recursive delay bound given with Theorem 1 should be used. Have in mind that such conclusions should be rather drawn for a specific scenario, as the gap between the homogeneous and heterogeneous delay bound varies for different path length and arrival data rates.

VI. Conclusion

Based on a recent approach in stochastic network calculus, we develop a bound on the probabilistic delay of a heterogeneous multi-hop wireless path with non-identically distributed channel gains. This bound represents first such available result, since an exact analytical result for the end-to-end performance of multi-hop wireless networks in terms of the underlying fading parameters does not exist yet, due to the inherent complexity of such systems and the numerous limitations of the available analytical methodologies. Hence, the discovered recursive bound provides a unique insight into the behaviour of probabilistic end-to-end delay in multi-hop wireless networks, as well as into the effect of fading on such networks' operation and performance. Apart from providing the analytical framework, we also validate the bound by means of simulation. The validation reveals in general a gap between the analytical bound and the simulation of about one order of magnitude, which also holds for very low delay violation probabilities. Such "conservative bound" may be of great importance for safety-critical applications demanding low violation probabilities. We further believe that this bound can be especially useful for network planning, admission control, routing and resource allocation. The independence of the derived bound of the underlying fading channel representation makes it applicable to various real-life systems, e.g. WirelessHART-based networks, machine-to-machine type of applications, etc. In particular, we are currently using the developed approach to analyze and optimize WirelessHART networks. Using the bound to develop channel-aware routing solutions by taking advantage of its recursive nature is the subject of future work.

References

[1] H. Al-Zubaidy, J. Liebeherr, and A. Burchard, "A (min, $\times$) Network Calculus for Multi-Hop Fading Channels," in *IEEE Infocom*, 2013.

[2] N. Bisnik and A. A. Abouzeid, "Queuing Network Models for Delay Analysis of Multihop Wireless Ad Hoc Networks," *Ad Hoc Networks*, vol. 7, no. 1, 2009.

[3] M. Xie and M. Haenggi, "Towards an End-to-End Delay Analysis of Wireless Multihop Networks," *Ad Hoc Networks*, vol. 7, no. 5, 2009.

[4] G. Gupta and N. Shroff, "Delay Analysis for Multi-Hop Wireless Networks," in *IEEE Infocom*, April 2009.

[5] S. Srinivasa and M. Haenggi, "Throughput-Delay-Reliability Tradeoffs in Multihop Networks with Random Access," in *48th Annual Allerton Conf. on Communication, Control and Computing*, Sept 2010.

[6] Y. Jiang, "A Basic Stochastic Network Calculus," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, Aug. 2006.

[7] K. Zheng, F. Liu, L. Lei, C. Lin, and Y. Jiang, "Stochastic Performance Analysis of a Wireless Finite-State Markov Channel," *IEEE Transactions Wireless Communications*, vol. 12, no. 2, February 2013.

[8] F. Ciucu, "Non-Asymptotic Capacity and Delay Analysis of Mobile Wireless Networks," *SIGMETRICS Perform. Eval. Rev.*, vol. 39, no. 1, Jun. 2011.

[9] F. Ciucu, O. Hohlfeld, and P. Hui, "Non-Asymptotic Throughput and Delay Distributions in Multi-Hop Wireless Networks," in *48th Annual Allerton Conf. on Communication, Control, and Computing*, Sept 2010.

[10] F. Ciucu, R. Khalili, Y. Jiang, L. Yang, and Y. Cui, "Towards a System Theoretic Approach to Wireless Network Capacity in Finite Time and Space," in *IEEE Infocom*, April 2014.

[11] M. Fidler, "A Network Calculus Approach to Probabilistic Quality of Service Analysis of Fading Channels," in *IEEE Globecom*, Nov 2006.

[12] R. McEliece and W. Stark, "Channels with block interference," *IEEE Trans. on Information Theory*, vol. 30 (1), 1984.

[13] M. Fidler, "An End-to-End Probabilistic Network Calculus with Moment Generating Functions," in *14th IEEE International Workshop on Quality of Service, 2006*, June 2006.

[14] D. Wu and R. Negi, "Effective Capacity: A Wireless Link Model for Support of Quality of Service," *IEEE Transactions Wireless Communications*, vol. 2, no. 4, July 2003.

[15] B. Davies, *Integral Transforms and Their Applications*. Springer-Verlag, 1978.