

# Estimating Traffic Correlations from Sampling and Active Network Probing

Amr Rizk, Zdravko Bozakov, and Markus Fidler  
 Institute of Communications Technology, Leibniz Universität Hannover  
 {amr.rizk, zdravko.bozakov, markus.fidler}@ikt.uni-hannover.de

**Abstract**—An extensive body of research deals with estimating the correlation and the Hurst parameter of Internet traffic traces. The significance of these statistics is due to their fundamental impact on network performance. The coverage of Internet traffic traces is, however, limited since acquiring such traces is challenging with respect to, e.g., confidentiality, logging speed, and storage capacity. In this work, we investigate how the correlation of Internet traffic can be reliably estimated from random traffic samples. These samples are observed either by passive monitoring within the network, or otherwise by active packet probes at end systems. We analyze random sampling processes with different inter-sample distributions and show how to obtain asymptotically unbiased estimates from these samples. We quantify the inherent limitations that are due to limited observations and explore the influence of various parameters, such as sampling intensity, network utilization, or Hurst parameter on the estimation accuracy. We design an active probing method which enables simple and lightweight traffic sampling without support from the network. We verify our approach in a controlled network environment and present comprehensive Internet measurements. We find that the correlation exhibits properties such as long range dependence as well as periodicities and that it differs significantly across Internet paths and observation times.

## I. INTRODUCTION

Traffic characteristics play a key role in planning and operation of packet data networks. As a consequence, in recent years network measurements have attracted considerable attention as a practical method for inferring traffic properties. The scope of such measurements varies from access networks to backbone networks or even across the Internet.

Numerous comprehensive measurement studies, based on recorded network traces, have revealed that aggregate Internet traffic possesses long memory correlations, so-called long range dependence (LRD) [7], [13], [20]. The impact of LRD on network performance was investigated in several works, e.g., [8], [16], [18], [24]. Networks fed with LRD traffic exhibit a fundamentally different behavior compared to systems fed with memoryless or Markovian traffic.

In practice continuous logging and evaluation of all relevant network events in large networks is typically not feasible due to efficiency, confidentiality, and cost factors. For example, with link speeds of 10 Gbps and more capturing traffic traces becomes increasingly difficult, as suitably large and fast storage systems are expensive. One main challenge is therefore, to extract the desired information from a subset of events, e.g., using a sampling procedure that yields consistent estimates of the target metric. In addition, ISPs rarely disclose traffic traces because of confidentiality issues such that traffic

characteristics can only be inferred from external observations. Further, a fundamental limitation of traffic traces is that these reflect traffic characteristics at only a *single* observation point.

In this work, we investigate the problem of estimating the correlation of Internet traffic given a limited set of random samples. The significance of these statistics is due to their fundamental impact on network performance [8], [13]. First, we consider passive sampling, i.e., capturing traffic samples at some directly accessible node, e.g., a router. Here, the main focus is on the choice of the sampling process and its properties. Further, for any practical realization passive sampling yields a finite sample size, which directly influences the accuracy of the results. Secondly, we consider active probing that is a technique, where external measurements of specific probe packets are used. The aim is to avoid any particular network support by exploiting, e.g., timing information that is imprinted on the probes by interaction with network traffic. The additional challenge of active compared to passive methods is to design probes that actually permit inferring the desired traffic characteristics, which in certain cases may even be impossible [15].

The ultimate result of this work is to enable the online estimation of traffic correlations along network paths without network support. To this end, we present methods for extracting LRD characteristics from sampled traffic. We derive the impact of sampling on the observed traffic correlations for different sampling strategies and show that sampling may distort observations. We develop methods that reverse these effects for a set of sampling processes. We quantify the accuracy of the observations under finite sampling durations, showing that the estimation error increases as  $\tau^{2-2H}$  with the autocovariance lag  $\tau$  and the LRD Hurst parameter  $H \in (0.5, 1)$ . We derive the impact of different sampling parameters on estimation accuracy and show a non-linear trade-off between sampling intensity and sampling duration. Finally, we design and evaluate a practical active probing method to estimate traffic correlations from external observations. We present practical testbed and Internet measurement results showing a complex covariance structure of Internet traffic that exhibits LRD as well as periodic behavior.

The paper is structured as follows: In the next section we present the state-of-the-art on LRD network traffic characteristics, sampling and active network probing. In Sect. III we derive our main results concerning traffic sampling and the accuracy of the estimated traffic parameters. In Sect. IV we

present and deploy an active probing method that uses packet probes to infer traffic correlations. Sect. V concludes the paper.

## II. RELATED WORK

In the following, we discuss related work on LRD traffic characteristics, sampling and network probing.

### A. LRD traffic characteristics

Comprehensive measurements in the 90s, e.g., [7], [13], [20] revealed that aggregate Internet traffic exhibits LRD and self-similarity phenomena, that can be described by the so-called Hurst parameter  $H$ . The aggregation of multiple traffic sources offers a possible explanation of these characteristics. It was shown in [31] that aggregating many on-off sources with heavy tailed on and off periods yields self-similar LRD traffic. This notion corresponds to file transfers from heavy tailed file size distributions as observed on storage systems [7], [34]. An experimental validation of the relation between self-similarity and heavy-tailed distributions is carried out in [14] on a large-scale experimental facility.

Given a stationary process  $Y(t)$ , LRD manifests itself in the slow decay of the autocovariance<sup>1</sup>  $c_Y(\tau)$  such that

$$c_Y(\tau) \sim \sigma_Y^2 \tau^{2H-2} \quad \text{for } \tau \rightarrow \infty, \quad (1)$$

where  $\sigma_Y^2$  is the variance of  $Y(0)$  and the Hurst parameter  $H \in (0.5, 1)$ . The sum of the autocovariance over all lags  $\tau$  diverges, i.e.,  $\sum_{\tau} c_Y(\tau) \rightarrow \infty$ .

In this work, we focus on the autocovariance structure of (1). Our goal is to infer (1) from traffic observations, respectively, to estimate the Hurst parameter  $H$  from the slope of  $c_Y(\tau)$  on a log-log scale. Numerous other methods exist for estimating the Hurst parameter from LRD and self-similar time series [3], [30], [33]. Due to limited space we refer the interested reader to our analysis of  $H$  estimation from sampled times series using a variance and a spectrum based technique in the extended technical report version of this paper [25].

### B. Sampling

Sampling is widely used to reduce the data processing and storage requirements as well as to circumvent problems, such as system inaccessibility and hardware access latency. A fundamental result often employed in the sampling context is known as PASTA, Poisson Arrivals see Time Averages [35]. PASTA states that the portion of Poisson arrivals that see a system in a certain state corresponds, in average, to the portion of time the system spends in that state.

Further, the authors of [17] establish general conditions, such that Arrivals See Time Averages (ASTA) holds, i.e., bias free estimates are not limited to Poisson sampling. In a recent work the authors of [2] coined the term NIMASTA, i.e. Non-intrusive Mixing Arrivals See Time Averages, in the context of network measurements using an argument on joint ergodicity.

<sup>1</sup>Throughout this work, we use the definition of autocovariance in the signal processing sense, i.e., for a stationary process  $Y(t)$  the autocovariance is defined as  $c_Y(\tau) := \mathbb{E}[Y(t)Y(t+\tau)] - \mathbb{E}[Y(t)]\mathbb{E}[Y(t+\tau)]$ . For brevity, we frequently use the term covariance to mean autocovariance.

The authors in [1] show that Poisson sampling, though bias free, does not guarantee minimum variance estimates.

A comparison of Poisson and periodic sampling was carried out in [27], [32]. Using the notion of asymptotic variance, [27] shows that either Poisson or periodic sampling can be superior depending on the a priori known autocovariance of the sampled process.

In [21] it is shown that for correlation lags tending to infinity, random sampling captures the long memory of the original processes, as long as the sampling distribution has a finite mean.

### C. Active network probing

The injection of test packets into a network for inferring network performance, i.e., active probing, has attracted considerable attention in recent years. End-to-end packet delays or inter packet times are metrics commonly used to estimate network characteristics such as the average available bandwidth or even to reconstruct cross-traffic statistics [12], [23], [29].

Cross traffic estimation of LRD traffic using active measurements was discussed, in [11], [22]. The authors of [11] carry out a numerical simulation to interpolate cross traffic from probes and predict future traffic from the LRD property. In [22] the authors derive and show simulation results for a deterministic probing scheme based on a multi-fractal wavelet traffic model. Essential to their estimation is the assumption that the queue does not empty between the individual packets of a packet probe. Our work differs significantly from [11], [22] as we examine *random* sampling distributions, show how to extract traffic correlations from distorted observations and characterize end-to-end paths.

Two important aspects concerning network probing are the measurement intrusiveness and the interaction of probes with the measured system. The first aspect is usually addressed by minimizing the probing rate while controlling the quality of the results. The second aspect is more involved, since the probes perturb the system leading to distorted observations. For example, measuring queueing delays of probes to determine the true queue length distribution is governed by a type of Heisenberg uncertainty [26], since the probes alter the queue length. The authors describe the impact of the probing intensity on the accuracy of the result using the notion of asymptotic variance. The effect is increased in case of LRD traffic, although not given in closed form, leading to higher uncertainty in the estimated waiting time [26].

## III. TRAFFIC SAMPLING AND PARAMETER ESTIMATION

In this section we derive our main results on traffic covariance estimation from sampled observations. Based on sampling properties we present rigorous traffic parameter estimation. Subsequently, we investigate the accuracy of the estimates under the practical constraint of finite sample sizes.

### A. Covariance of sampled processes

We define a sampling model comprising of three stationary discrete time processes: a traffic increment process  $Y(t)$ , a

TABLE I  
PARAMETRIZATION OF SAMPLING DISTRIBUTIONS AND TRAFFIC PARAMETER ESTIMATION

	inter-sample distribution $f(\tau)$	autocovariance $c_A(\tau) =$ $E[A(t)A(t+\tau)] - \mu_A^2$ for $\tau > 0$	reconstructed traffic covariance $c_Y(\tau)$	remarks
Geometric	$p(1-p)^{\tau-1}$	0	$\frac{c_W(\tau)}{\mu_A^2}$	$\mu_A = p$
Periodic	$\delta(\tau - \Delta)$	$\begin{cases} 1/\Delta - 1/\Delta^2 & \text{for } \tau = k\Delta, k \in \mathbb{N}_0 \\ -1/\Delta^2 & \text{otherwise} \end{cases}$	$\frac{c_W(\tau) - \mu_A \mu_Y^2 (1 - \mu_A)}{\mu_A}$	$c_Y(\tau)$ at $\tau = k\Delta,$ $\mu_A = 1/\Delta$
Gamma	$\frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta\tau}$	$-\mu_A^2 e^{-4(\mu_A)\tau}$	$\frac{c_W(\tau) + \mu_A^2 \mu_Y^2 e^{-4(\mu_A)\tau}}{\mu_A^2 (1 - e^{-4(\mu_A)\tau})}$	for $\alpha = 2, \mu_A = \frac{\beta}{\alpha}$
Uniform	$1/b$ for $0 \leq \tau \leq b$	$\mu_A^2 (\frac{1}{2} e^{\frac{1}{2}(\mu_A)\tau} - 1)$	$\frac{2c_W(\tau) - \mu_A^2 \mu_Y^2 (e^{\frac{1}{2}(\mu_A)\tau} - 2)}{\mu_A^2 e^{\frac{1}{2}(\mu_A)\tau} - 2}$	$c_Y(\tau)$ for $\tau \leq b,$ $\mu_A = 2/b$

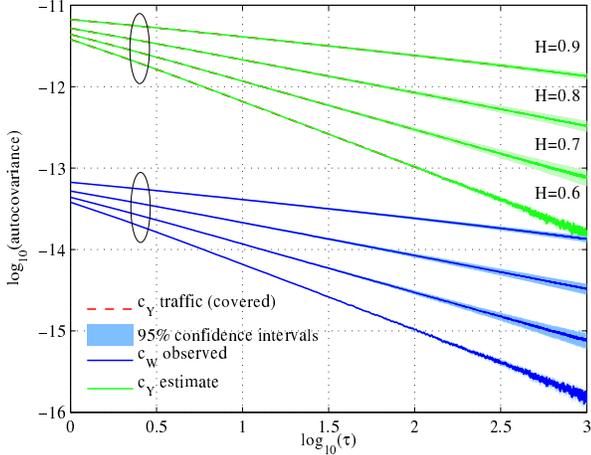


Fig. 1. Autocovariance of LRD traffic processes under geometric sampling. The observed “ $c_W(\tau)$ ” maintains the autocovariance structure of the traffic process. The covariance of the original process “ $c_Y(\tau)$  (traffic)” is exactly covered by the reconstructed “ $c_Y(\tau)$  (estimate)”.

sampling process  $A(t)$ , and an observed process  $W(t)$  for  $t \in \mathbb{N}_0$ . We assume statistical independence of  $A(t)$  and  $Y(t)$ . Our focus lies on the estimation of the covariance of  $Y(t)$  that is characterized by LRD. While the LRD process may be in continuous time, we regard its increments on a fixed time slot basis, and hence the discretization of  $Y(t)$ .

The sampling process  $A(t)$  is a point process taking the value of one whenever a sample is taken, and zero otherwise, i.e.,  $A(t)$  is a Kronecker delta train, where a Kronecker delta is defined as  $\delta(n) = 1$  for  $n = 0$  and zero otherwise. The process has independent and identically distributed (iid) inter-sample times drawn from a given probability distribution  $F$ . The inter-sample time is the time between two consecutive Kronecker deltas. The sampling intensity, i.e., the mean rate of the sampling process of  $A(t)$ , is  $E[A(t)] = \mu_A$  for all  $t$ , with  $0 \leq \mu_A \leq 1$ . Throughout this work we use  $\mu(\cdot)$  to denote the expected value  $E[(\cdot)]$ .

We base our analysis on the observed stochastic process  $W(t)$ , generated by random samples  $A(t)$  of the increment process  $Y(t)$ , with

$$W(t) = A(t)Y(t). \quad (2)$$

We aim to infer properties of the traffic process  $Y(t)$  from the observation process  $W(t)$ . In particular, we are interested

in sampling distributions  $F$  that deliver accurate estimates of the correlations of the LRD traffic process  $Y(t)$  and the associated Hurst parameter  $H$ . Extracting the autocovariance of the process  $Y(t)$ , i.e.,  $c_Y(\tau)$  from the observed  $c_W(\tau)$  is generally not a straightforward task. The following lemma reveals the impact of sampling on the autocovariance of the observed process. The proof of Lem. 1 is a variation of standard technique in stochastic and is given in the technical report [25].

*Lemma 1:* Given the stationary and independent stochastic processes  $A(t)$  and  $Y(t)$  and let  $W(t) = A(t)Y(t)$ . The covariance of  $W(t)$  can be decomposed into

$$c_W(\tau) = (c_A(\tau) + \mu_A^2) c_Y(\tau) + c_A(\tau) \mu_Y^2.$$

Lem. 1 clearly shows the impact of the sampling process on the observed covariance. In particular, the choice of the inter-sample distribution influences  $c_W(\tau)$  through  $\mu_A$  and  $c_A(\tau)$ , i.e., both the sampling intensity and the sampling covariance influence the observation.

In this work we investigate four inter-sample distributions: geometric (memoryless), periodic, Gamma, and uniform. For each distribution we show how to recover the covariance of the LRD process  $c_Y(\tau)$  from the observed  $c_W(\tau)$  using the covariance  $c_A(\tau)$ . To this end, we derive the covariance of the sampling process  $c_A(\tau) = E[A(t)A(t+\tau)] - \mu_A^2$ . We use the probability mass function  $f(\tau)$  of the inter sample times to calculate the  $n$ -fold self-convolution  $f^{(*n)}(\tau)$ . We then calculate the autocorrelation  $E[A(t)A(t+\tau)] = \mu_A \sum_{n=1}^{\infty} f^{(*n)}(\tau)$  as given in [6], Eq. (4.6.1). We exploit the property that  $f^{(*n)}$  is a power series for the considered distributions and that its sum converges. We provide an elaborate derivation of the autocovariance in the technical report [25]. In the last step we insert  $c_A(\tau)$  into Lem. 1 and solve for  $c_Y(\tau)$ .

Tab. I summarizes the expressions used to reconstruct  $c_Y(\tau)$  given specific inter-sample distribution parameters and corresponding  $c_A(\tau)$ . First, we consider the geometric inter-sample distribution, i.e., a Bernoulli sampling process. The independence of the increments implies that  $c_A(\tau) = 0$  for  $\tau > 0$ . From Lem. 1, the observations  $W(t)$  have autocovariance

$$c_W(\tau) = c_Y(\tau) \mu_A^2. \quad (3)$$

This indicates that sampling processes with uncorrelated increments preserve the autocovariance structure of  $Y(t)$ .

Next, we consider periodic sampling, where  $A(t)$  is modeled as a comb of Kronecker deltas with sampling period  $\Delta$ . The mean intensity of the sampling process is  $\mu_A = 1/\Delta$ . We can recover  $c_Y(\tau)$  at  $\tau = k\Delta$ , however, the mean rate  $\mu_Y$  of the traffic process must be known. Due to the rigid structure of periodic sampling it is known that the associated mean rate estimator  $\mu_W/\mu_A$  is not unbiased [2], e.g., the sampling period may coincide with periodicities in the original process.

Finally, Tab. I provides expressions for reconstructing  $c_Y(\tau)$  after Gamma and uniform sampling. For mathematical tractability, here we use continuous time for the derivation of the autocorrelation of  $A(t)$ . Note that the discretization error diminishes for autocorrelation lags much larger than the discretization time slot. In case of Gamma sampling, the ability to estimate  $c_Y(\tau)$  is not limited to the exemplary  $\alpha = 2$  given in Tab. I. Lem. 1 can be used to estimate  $c_Y(\tau)$  for arbitrary Gamma sampling processes as long as the autocovariance  $c_A(\tau)$  is computable. We provide results for Gamma sampling with  $\alpha = 4$  in the technical report [25].

Figures 1 and 2 illustrate autocovariance estimates derived from observations  $W(t)$ , that are obtained by sampling LRD traffic with autocovariance  $c_Y(\tau) \sim \sigma_Y^2 \tau^{2H-2}$  and  $H \in [0.6, 0.9]$ .<sup>2</sup> We use geometric, periodic, Gamma, and uniform inter-sample time distributions and set  $\mu_A = 0.1$ . In all cases the reconstructed autocovariance denoted “ $c_Y(\text{estimate})$ ” exactly covers the original traffic autocovariance “ $c_Y(\text{traffic})$ ”.

Geometric sampling in Fig. 1 preserves the linear covariance structure of  $c_Y(\tau)$ . The observed  $c_W(\tau)$  is vertically shifted by  $\log(\mu_A^2)$  w.r.t. the original  $c_Y(\tau)$ . The Hurst parameter  $H$  can be inferred directly from the slope of  $c_W(\tau)$ .

For the remaining considered distributions shown in Fig. 2, the observations  $c_W(\tau)$  are distorted. However, using Lem. 1 we recover the original covariance  $c_Y(\tau)$ . Using the expressions from Tab. I we reconstruct “ $c_Y(\text{estimate})$ ” which lies on top of the original autocovariance “ $c_Y(\text{traffic})$ ”.

In the following we discuss advantages and disadvantages of the presented sampling distributions. Periodic and uniform sampling are practically convenient as the inter-sample times

cannot become arbitrarily large due to the finite support of the inter-sample distribution. Moreover, periodic sampling is easy to implement.

However, it is important to point out that periodic sampling yields misleading results if the sampling period coincides with periodicities in the target process. In addition, periodic, Gamma as well as uniform sampling require a reconstruction step to estimate the covariance  $c_Y(\tau)$  from observations as shown above. To this end, an estimate of the mean rate of  $Y(t)$  is required.

Memoryless sampling is proposed by the IETF as a network probing scheme [19]. In contrast to periodic, Gamma and uniform sampling, a major advantage of geometric sampling, i.e., memoryless, is that the covariance structure of  $c_Y(\tau)$  is preserved in the observations as given in (3). In the following we continue the analysis with geometric sampling because of its advantages discussed above.

### B. Impact of finite sample sizes

Next, we examine the accuracy of the derived estimates for finite sample sizes which is vital for any practical realization. The use of finite sample sizes relaxes the assumption of stationarity to piece-wise stationarity for the duration of a measurement. We determine the impact of sampling parameters, e.g., sampling duration or intensity, on the observations. Moreover, we evaluate the accuracy of the deployed statistical estimators. Finally, we recover the results from Sect. III-A in the limit for infinite sampling durations.

We investigate sample autocovariances marked by  $\tilde{c}_{(\cdot)}$  as estimators of the population autocovariances  $c_{(\cdot)}$ . In addition, we consider the sample means  $\tilde{\mu}_{(\cdot)}$  as estimators of the population means  $\mu_{(\cdot)}$ . To better understand the impact of finite sample sizes on the observations and the covariance estimates we examine the individual effects of the sample covariances involved in a step by step manner.

While geometric sampling is appealing since it's autocovariance  $c_A(\tau) = 0$  for  $\tau > 0$ , it loses this property for finite sampling duration  $T$ , where  $T$  is the length of the time-slotted sampling process  $A(t)$  in slots.

In the following we focus on three aspects. First in subsection III-B1, we derive the impact of finite sample sizes on the

<sup>2</sup>Synthetic traces of length  $2.5 \times 10^8$  time slots were used for the simulation which was repeated 25 times for each considered  $H$ .

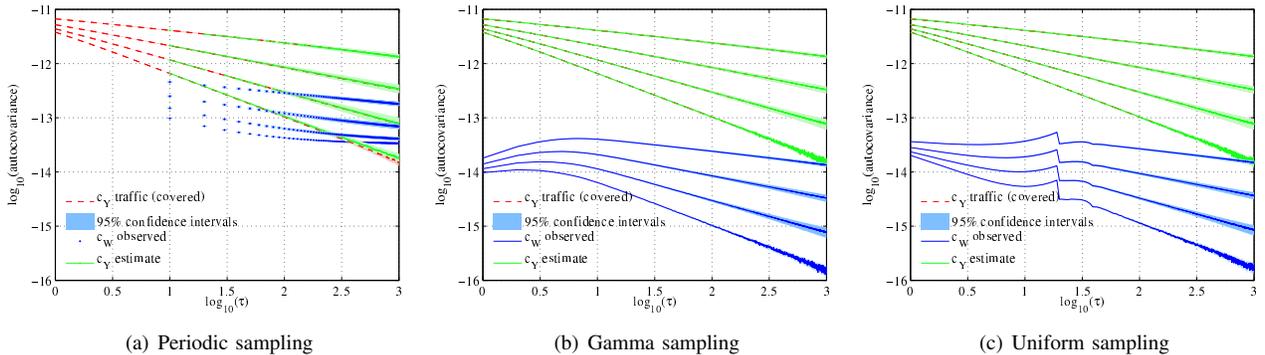


Fig. 2. Autocovariance of the LRD process under different sampling strategies. Note that “ $c_Y(\tau)$  (traffic)” is covered by the “ $c_Y(\tau)$  (estimate)”.

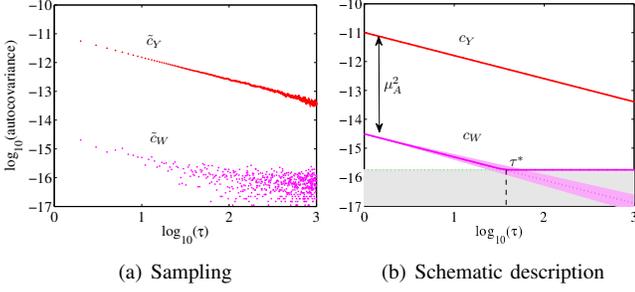


Fig. 3. Noisy observations due to finite sampling. Noise floor (shaded area) in Sect. III-B1. Noise cone in Sect. III-B2.

observability of the covariance of sampled traffic. The second aspect is the impact of the sample covariance  $\tilde{c}_A(\tau)$  and its influence on the estimation error. This is handled in subsection III-B2. The third aspect is the impact of finite sample sizes on the bias of the covariance estimators given in subsection III-B3.

1) **Observation limit:** In this subsection, we do not consider deviations of sample statistics from respective population measures. We relax this assumption in the following subsections. Fig. 3(a) depicts the sample autocovariance  $\tilde{c}_Y(\tau)$  of an LRD traffic trace  $Y(t)$ , and the corresponding autocovariance of geometrically sampled observations  $\tilde{c}_W(\tau)$  with a limited sample size  $T$ . Evidently,  $\tilde{c}_W(\tau)$  is not just a shifted version of  $\tilde{c}_Y(\tau)$  but superposed and distorted with observation “noise” for increasing lags  $\tau$ .

We seek a range of lags  $\tau \in [0, \tau^*]$  in which the covariance of the sampled process can be observed without significant distortion. Based on a standard technique [3] we compare the covariance of the observed process to the covariance of geometrically sampled iid Gaussian sequences to obtain  $\tau^*$  up to which both covariances are significantly different.

Fig. 3(b) schematically depicts  $\tau^*$  as the intersection of  $c_W(\tau)$  from (3) and the shaded 0.95 confidence interval for geometrically sampled finite Gaussian iid sequences with mean  $\mu_Y$  and variance  $\sigma_Y^2$ . For  $T \gg \tau$  we find that this confidence interval is given by  $2(\sigma_A^2 \mu_Y^2 + \mu_A \sigma_Y^2) \sqrt{(4\mu_A^2 \mu_Y^2)/(\sigma_A^2 \mu_Y^2 + \mu_A \sigma_Y^2) + 1}/\sqrt{T}$ . The calculation relies on the central limit theorem and is given in detail in the technical report [25]. In Fig. 3(b) we denote this confidence interval as noise floor.

We calculate  $\tau^*$  for LRD traffic with covariance  $c_Y(\tau) = K\sigma_Y^2 \tau^{2H-2}$ , with constant  $K$ , as

$$\tau^* = \left[ \frac{K\sigma_Y^2 \mu_A \sqrt{T}}{2(\sigma_A^2 \mu_Y^2 + \mu_A \sigma_Y^2) \sqrt{\frac{4\mu_A^2 \mu_Y^2}{\sigma_A^2 \mu_Y^2 + \mu_A \sigma_Y^2} + 1}} \right]^{\frac{1}{2-2H}}.$$

It is obvious that stronger LRD, i.e., higher  $H$ , is observed better. Clearly, for an infinite sample size  $T \rightarrow \infty$ , the observable range goes to infinity  $\tau^* \rightarrow \infty$ . Fig. 3(a) shows that in practice it is important to consider this range to ensure that the results are not significantly distorted.

2) **Estimation accuracy:** Next, we evaluate the impact of the finite sample size on the sample covariance  $\tilde{c}_A(\tau)$ . We

analyze the influence of  $\tilde{c}_A(\tau)$  on the observation  $\tilde{c}_W(\tau)$  and of estimates of  $c_Y(\tau)$  obtained thereof. For ease of exposition, we assume  $\tilde{c}_Y(\tau) = c_Y(\tau)$ ,  $\tilde{\mu}_A = \mu_A$  and  $\tilde{\mu}_Y = \mu_Y$ , i.e., in this subsection we restrict our analysis to the deviation of  $\tilde{c}_A(\tau)$  from  $c_A(\tau)$ .

We assume  $T \gg \tau$  and use the central limit theorem to approximate the distribution of the sample autocovariance  $\tilde{c}_A(\tau)$  by a Gaussian distribution with standard deviation  $\sigma_A \sqrt{\sigma_A^2 + 4\mu_A^2/\sqrt{T-\tau}}$ , with  $\sigma_A^2 = \mu_A - \mu_A^2$  from the geometric distribution. We calculate the 0.95 confidence interval  $c_A^{95} \approx \pm 2\sigma_A \sqrt{\sigma_A^2 + 4\mu_A^2/\sqrt{T-\tau}}$  for the mean sample autocovariance<sup>3</sup>. The derivation can be found in the appendix of the technical report [25].

With help of  $c_A^{95}$  we investigate the impact of the variations of  $\tilde{c}_A(\tau)$  on the observation  $\tilde{c}_W(\tau)$ . First, we use  $c_A^{95}$  to calculate a confidence interval for  $\tilde{c}_W(\tau)$  as  $c_W^{95}(\tau) \approx \pm c_A^{95}(c_Y(\tau) + \mu_Y^2)$ . We schematically depict  $c_W^{95}(\tau)$  as noise cone in Fig. 3(b).

Next, in reference to (3) we consider the estimator  $\tilde{c}_W(\tau)/\mu_A^2$  for estimating  $c_Y(\tau)$ . We analyze the impact of the variations of  $\tilde{c}_A(\tau)$  on this estimator. We calculate the confidence interval  $c_Y^{95}(\tau)$  for this estimator as  $c_Y^{95}(\tau) \approx \pm c_A^{95}(c_Y(\tau) + \mu_Y^2)/\mu_A^2$ . Finally, we obtain the following relative error

$$\varepsilon_Y^{rel}(\tau) = \frac{|c_Y^{95}(\tau)|}{c_Y(\tau)} \approx \frac{2\sigma_A \sqrt{\sigma_A^2 + 4\mu_A^2}}{\sqrt{T-\tau}\mu_A^2} \left( 1 + \frac{\mu_Y^2}{c_Y(\tau)} \right). \quad (4)$$

From (4) we observe that the estimation error introduced through  $\tilde{c}_A(\tau)$  decays with increasing sampling duration  $T$  or with increasing sampling intensity  $\mu_A$ . For small (practical) sampling intensities, e.g.,  $\mu_A \leq 0.1$ , we find a nonlinear trade-off between sample intensity  $\mu_A$  and sampling duration  $T$ . Using  $\sigma_A^2 = \mu_A - \mu_A^2$  from the geometric sampling distribution the prefactor in (4) can be approximated as  $1/\sqrt{T}\mu_A$  for  $T \gg \tau$ . This result enables the important conclusion that for finite sample sizes sampling intensity has a stronger impact on accuracy than sampling duration.

Next, we examine the influence of the parameter  $H$  on (4) for large lags  $\tau$ . For increasing  $\tau$ ,  $c_Y(\tau)$  decreases, such that when  $c_Y(\tau) \ll \mu_Y^2$ , the relative estimation error (4) becomes

$$\varepsilon_Y^{rel}(\tau) \approx \frac{2\sigma_A \sqrt{\sigma_A^2 + 4\mu_A^2} \mu_Y^2}{\sqrt{T-\tau}\mu_A^2 \sigma_Y^2} \tau^{2-2H}. \quad (5)$$

The relative estimation error  $\varepsilon_Y^{rel}(\tau)$  increases with the lag  $\tau$  depending on  $H \in (0.5, 1)$ . For LRD traffic which exhibits large  $H$ , the estimation error increases slower in  $\tau$  compared to traffic with a small parameter  $H$ .

We depict  $\varepsilon_Y^{rel}(\tau)$  in Fig. 4. To this end, we used 100 generated LRD traffic traces with  $T = 2 \times 10^8$  time slots. The figure includes auxiliary lines with a slope of  $2-2H$ . It is evident, that the estimation error evolves with  $\tau$  as in (5).

In addition, we calculate the needed sampling duration  $T$  to achieve constant  $\varepsilon_Y^{rel}$  for a given lag  $\tau$ , and fixed  $\mu_A, \mu_Y$  and

<sup>3</sup>We use the relation  $\approx$  to denote the approximation, here due to the Gaussian distribution approximation.

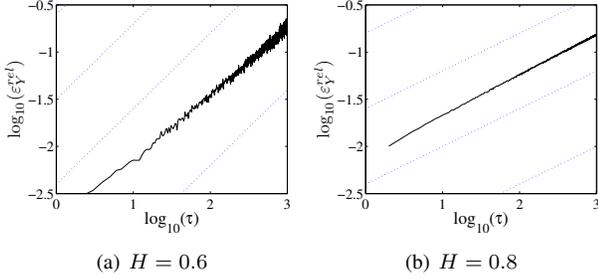


Fig. 4. Estimation error under finite sampling depends on  $H$ .

$\sigma_Y$ . We find from (5) that the sampling duration has to increase as  $T \sim \max\{\tau^{4-4H}, \tau\}$ , which again reveals the impact of  $H$ . Specifically, for  $H < 0.75$  the sampling duration has to increase faster than linearly with  $\tau$  to achieve constant  $\epsilon_Y^{rel}$ .

3) **Bias of autocovariance estimators:** Next we investigate the accuracy of the deployed statistical estimators. The impact of the finite sample size carries forward to the computation of the autocovariance of  $Y(t)$ . First we consider the case where we directly observe  $Y(t)$  for a finite duration  $T$ . We consider the autocovariance estimator  $\tilde{c}_Y(\tau) = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} (y(t) - \tilde{\mu}_{Y_0})(y(t+\tau) - \tilde{\mu}_{Y_\tau})$  with  $\tilde{\mu}_{Y_i} = \frac{1}{(T-\tau)} \sum_{t=1}^{T-\tau} y(t+i)$ . An estimator of the autocovariance is unbiased iff  $E[\tilde{c}_Y(\tau)] = c_Y(\tau)$ . To inspect the bias of  $\tilde{c}_Y(t)$ , we calculate its expected value and find

$$E[\tilde{c}_Y(\tau)] \approx c_Y(\tau) - \frac{\sigma_Y^2}{(T-\tau)^{2-2H}}. \quad (6)$$

The derivation of (6) is given in the appendix of the technical report [25].

From (6) we conclude that the autocovariance estimator  $\tilde{c}_Y(\tau)$  is asymptotically unbiased for  $T \rightarrow \infty$  and  $T \gg \tau$ . The maximum lag, up to which the autocovariance is estimated, must be chosen carefully, such that the bias in (6) becomes negligible. However, the bias depends on  $H$  such that higher  $H$  require larger  $T$  to retain a negligible bias.

After considering the entire process  $Y(t)$  we now investigate the bias of the autocovariance estimator when applied to  $W(t)$  as observed by sampling with finite duration  $T$ . We calculate the expected value of the estimated autocovariance

$$E[\tilde{c}_W(\tau)] \approx c_W(\tau) - \frac{c_W(0)}{T-\tau} - \frac{2}{(T-\tau)^2} \sum_{t=1}^{T-\tau-1} (T-\tau-t)c_W(t). \quad (7)$$

The derivation of (7) is given in the appendix of the technical report [25]. The bias in (7) goes to zero for  $T \rightarrow \infty$  and  $T \gg \tau$ .

In the remainder of this section we provide brief conclusions that highlight our main findings. We presented a framework for extracting the traffic autocovariance from observed samples. From our evaluation of the sampling distributions we conclude, that the covariance observed under geometric sampling does not exhibit any distortions. This property greatly simplifies the reconstruction of the covariance of the original process  $Y(t)$ , as no additional parameters, such as  $\mu_Y$ , must be estimated. Hence, for geometric sampling with sufficiently large  $T$  we

use  $\tilde{c}_W(\tau)/\mu_A^2$  as an estimator of the traffic autocovariance. From the evaluation of the estimator we find two major aspects that limit the observability for finite sampling sizes. First, finite sampling size yields a computable noise range which may obscure the true covariance structure. Secondly, the bias for covariance estimators depends on the Hurst parameter, such that longer measurements must be conducted for traffic exhibiting strong LRD.

Nevertheless, finite sampling effects disappear in the limit for large sampling durations. Moreover, we found that increasing the probing intensity improves estimation results more quickly than increasing the sampling duration.

#### IV. ACTIVE PROBING

So far, we focused on the estimation of traffic correlations using passive sampling. In large multi-provider networks like the Internet, service providers often do not provide such network traces, e.g., for reasons of competition. The estimation of traffic correlations, therefore, must rely on inferring samples of the Internet traffic from network metrics that can be easily observed at end systems, e.g., by active probes. Moreover, passive sampling is a priori limited to single links. In case of network paths, where the correlations of the end-to-end service involve multiple nodes and links, end-to-end measurements may be the only viable option. We present an active probing method that enables users to characterize end-to-end paths, with minimal effort and without administrative support from the network under observation.

In this section, we address the fundamental problem of inferring the correlation of LRD traffic using active probes. We propose a new active probing method which collects traffic samples by detecting router busy periods. The observations are used to estimate the covariance of the end-to-end service. Subsequently, we estimate the corresponding Hurst parameter. In the extended technical report [25] we show an alternative approach for estimating traffic correlations that is based on capturing traffic intensities using packet pair probes. Compared to packet pairs, the approach described in the following uses less probing traffic and can be stringently formulated for multi-node networks. In the sequel, we describe our probing methodology and discuss traffic correlation estimation for both the single and multi-node cases. We then show testbed measurements to demonstrate the feasibility of our method. Finally, we present a set of Internet measurement results showing end-to-end correlations of entire network paths.

##### A. Probing Methodology

To extract an estimate of the cross traffic autocovariance, we propose an approach which uses the delays of single packet probes to detect busy periods at a router, and hence samples the link utilization at the router egress. For the remainder of this work, cross traffic denotes any traffic sharing resources with the probing traffic.

We make the general assumption that packet scheduling is non-preemptive. Hence, whenever a router is busy transmitting a packet, the delay  $d_p$  experienced by an arriving packet will

be greater than the minimal delay  $d_{min}$  experienced when the router is idle. Consequently, we can sample cross traffic increments at the router egress, by injecting probe packets and analyzing their delays. For each probe, we measure the one way delay  $d_p = t_r - t_s$ , using the send and receive times  $t_s$  and  $t_r$ , respectively. To determine if the router was busy, we check whether the observed delay  $d_p$  is greater than the minimum network delay  $d_{min}$ . As a result, each probe yields a sample of the egress link state at time  $t$ , and the observed process can be constructed as

$$W(t) = \begin{cases} 1 & \text{if } d_p > d_{min} \text{ and } A(t) = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

It is known [8], [9], that the covariance structure of LRD traffic is preserved at the output of a queue or a traffic shaper, such that  $W(t)$  permits observing the covariance of the cross traffic. We assume that the perturbation of the observed traffic due to probe size and probing rate is negligible, since the probing rates used are typically less than one per mill of the capacity. Furthermore, as we can assume that dropped probes are due to a busy router, we account for lost probing packets by setting  $W(t) = 1$  for all dropped probes of  $A(t)$ .

### B. Measuring LRD in single- and multi-node scenarios

We now show that the observed process  $W_N(t)$  at the egress of an  $N$  node path with LRD cross traffic also exhibits LRD behavior. Moreover, for cross traffic characterized by different Hurst parameters, we show that the largest Hurst parameter dominates the covariance of the observed process  $W_N(t)$ . These results are in agreement with [9] which shows that the largest Hurst parameter dominates end-to-end performance.

Consider an  $N$  node topology with independent LRD cross traffic as in Fig. 5. We describe the busy state of each node using the processes  $Y_i(t)$  for node  $i \in [1, N]$ . Hence,  $Y_i(t) = 1$  if node  $i$  is busy at time  $t$  and  $Y_i(t) = 0$  otherwise. Note that the covariance  $c_{Y_i}(\tau) \sim \tau^{2H_i-2}$  measured at egress of node  $i$  has the same LRD property as the cross traffic input at the node [8], [9]. Next, consider an active probe that is injected into the path. After subtracting the minimum end-to-end delay  $d_{min}$  the observer at the egress of the path will measure a positive delay only if any of the routers were busy when the probe arrived at the respective router. Otherwise, the probe delay will equal zero. Hence,  $W_N(t)$  is the logical OR operation of the individual processes  $Y_i(t)$  for  $i \in [1, N]$ . Since  $Y_i(t)$  and  $W_i(t) \in \{0, 1\}$ , we straightforwardly find  $W_i(t)$  at the egress of node  $i$  as

$$W_i(t) = \begin{cases} Y_1(t) & , i = 1 \\ W_{i-1}(t) + Y_i(t) - W_{i-1}(t)Y_i(t) & , i \in [2, N]. \end{cases} \quad (9)$$

We denote  $E[Y_i(t)] = \mu_{Y_i}$  and  $E[W_i(t)] = \mu_{W_i}$  for  $i \in [1, N]$ . First, we illustrate (9) using a two node example and two independent LRD processes  $Y_1(t)$ ,  $Y_2(t)$ . The observed process at the egress of node 2 is  $W_2(t) = 1$  if  $Y_1(t) = 1$  OR  $Y_2(t) = 1$  and  $W_2(t) = 0$  otherwise, such that we deduce

$$W_2(t) = Y_1(t) + Y_2(t) - Y_1(t)Y_2(t).$$

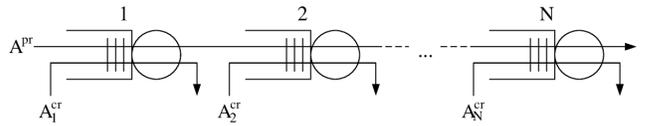


Fig. 5.  $N$  node topology with probing traffic  $A^{pr}$  and LRD cross traffic  $A_i^{cr}$  for  $i \in [1, N]$ .

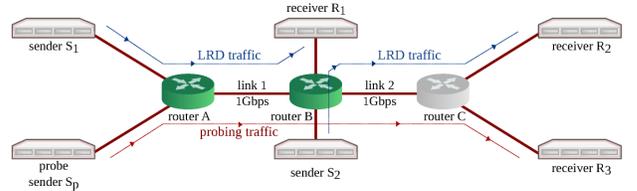


Fig. 6. Experimental setup: Emulab testbed

We derive the observed covariance  $c_{W_2}(\tau)$  of  $W_2(t)$  after some reordering as

$$c_{W_2}(\tau) = c_{Y_1}(\tau)c_{Y_2}(\tau) + c_{Y_1}(\tau)(1 - \mu_{Y_2})^2 + c_{Y_2}(\tau)(1 - \mu_{Y_1})^2.$$

The equation above directly shows that for large  $\tau$  the covariance  $c_{W_2}(\tau)$  is dominated by  $c_{Y_i}(\tau)$  with the largest Hurst parameter, i.e., slowest decay. The covariance of the  $N$ -node end-to-end observations  $c_{W_N}(\tau)$  is obtained using the recursion formula (9) as

$$c_{W_i}(\tau) = c_{W_{i-1}}(\tau)c_{Y_i}(\tau) + c_{W_{i-1}}(\tau)(1 - \mu_{Y_i})^2 + c_{Y_i}(\tau)(1 - \mu_{W_{i-1}})^2. \quad (10)$$

Using recursive substitution it can be shown that the covariance of the end-to-end observations  $c_{W_N}(\tau)$  is dominated by  $c_{Y_i}(\tau)$  with the largest Hurst parameter for  $i \in [1, N]$ .

### C. Probing Software and Experimental Setup

We developed a probing tool  $H$ -probe, available at [4], that implements the method above and the statistical analysis discussed in Sect. III-B. Our method from Sect. IV-A does not require assumptions about the content and type of packets that are used as probes.  $H$ -probe injects ICMP echo requests from the sender to an arbitrary receiver and captures the associated round trip times using *libpcap*. We made the practical choice of using ICMP round trip times to avoid the need for clock synchronization. Further, ICMP allows probing the path to any network host without the need for a receiver software. We note, that it is straight-forward to extend  $H$ -probe to other protocols such as UDP. Hence,  $H$ -probe infers the covariance structure of the round trip service of network paths using online measurements.

In the following we present results obtained using this software package. Fig. 6 depicts the experimental setup in our Emulab-based testbed<sup>4</sup>. The topology comprises two relevant links, denoted link 1 and 2. Two traffic senders  $S_i$ ,  $i \in [1, 2]$  transmit LRD cross traffic traces with defined Hurst parameter

<sup>4</sup>We use nodes with Supermicro X8DTU server mainboards with 2.2Ghz Intel E5520 Xeon processors, quad port Intel 82576EB Gigabit Ethernet Controllers, and Ubuntu 10.04 LTS with kernel 2.6.32-24. All links have a capacity of  $C = 1$ Gbps.

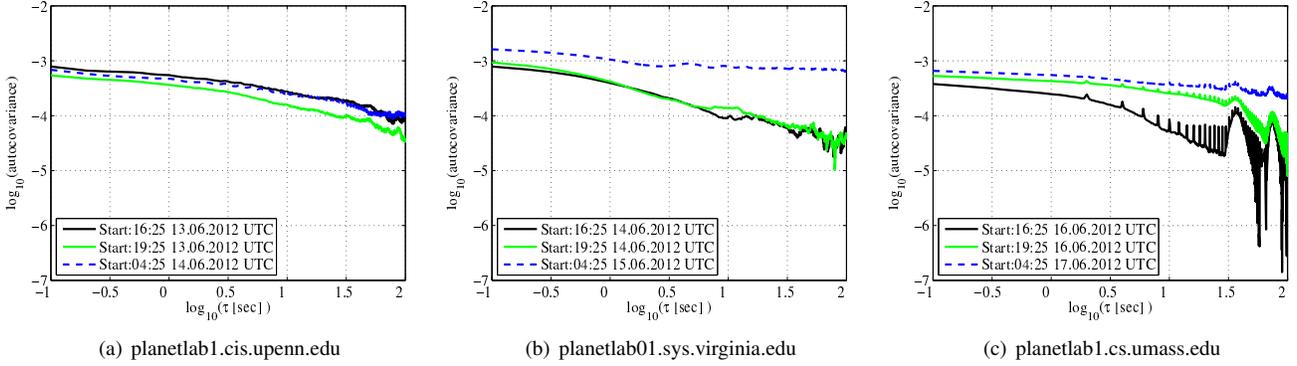


Fig. 8. End-to-end covariance estimates from Internet measurements. The covariance structure varies across different paths and for different times. For some targets we observe distinct periodicities on different timescales.

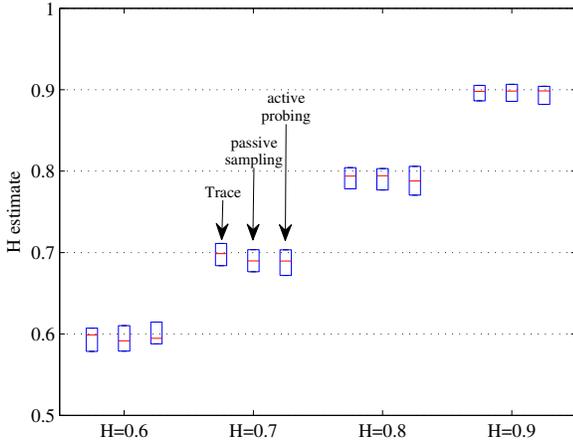


Fig. 7. Hurst parameter estimates from (a) offline trace analysis, (b) offline trace sampling and (c) active probing in the Emulab testbed.

$H$  to the receivers  $R_i$ . The traces were synthesized by superposition of  $10^5$  heavy tailed on-off sources with tail index  $\alpha$ . The relation between  $H$  and the tail index  $\alpha$  is given in [34]. We set the mean rate of the traffic at *each* sender to 50 Mbps, with a constant packet size of 1500 Byte.

We use geometrically distributed inter-sample times with  $p = 0.1$  and slot length  $\delta = 1$ ms. For each measurement we send  $10^6$  probes with a mean probing rate of 100 packets per second (corresponding to  $\sim 70$  kbps) from the probe sender  $S_p$  to the receiver  $R_3$ . We fix the measurement duration to 3 hours that is a time-scale over which piecewise stationarity of traffic processes has been observed, e.g. in [10], [28]. The chosen probing parameters introduce a light load on the nodes and simultaneously yield viable results. We highlight the non-linear tradeoff between probing duration and intensity in Sect. III-B. We use the same parameters for the Internet measurements in Sect. IV-E. To deal with non-queueing induced jitter in routers, which we assume to be light tailed,  $H$ -probe substitutes  $d_{min}$  from (8) by the strictly larger average  $E[d]$  to reduce the measurement noise. While this heuristic conceals small bursts, we note that the long tail of the burst length distribution, which establishes the LRD property of the traffic [34], remains unaffected.

TABLE II  
EXEMPLARY HURST PARAMETER ESTIMATES IN A 2 NODE SCENARIO.

	estimated $H$ on run #				
	1	2	3	4	5
$\{H_1 = 0.6, H_2 = 0.9\}$	0.87	0.89	0.89	0.90	0.90
$\{H_1 = 0.9, H_2 = 0.6\}$	0.87	0.88	0.88	0.90	0.90
$\{H_1 = 0.6, H_2 = 0.6\}$	0.59	0.62	0.64	0.63	0.63
$\{H_1 = 0.9, H_2 = 0.9\}$	0.92	0.92	0.89	0.92	0.89

#### D. Testbed measurements

We deploy  $H$ -probe in our Emulab testbed, in order to verify its functionality in a controlled environment. We use synthetic traces to be able to repeat the experiment for statistical validity. First, we inject synthetic LRD traffic with  $H \in [0.6, 0.9]$  on link 1 and collect  $10^6$  samples using our software. In order to compare the covariance of the traffic traces to the measurement results we do not inject traffic on the return path. Each experiment is repeated 25 times. We compare the covariance of the full traffic traces calculated offline (denoted trace) to the covariance extracted offline from a sampled process (denoted passive sampling) as well as from probes using  $H$ -probe (denoted active probing). To this end, we estimate the Hurst parameter using a least square regression of the estimated covariance on lags  $\tau \in [1, 10^3]$ . The lag range for the regression as well as the probing process parameters are chosen according to the constraints in Sect. III-B. We show boxplots of the corresponding Hurst parameters in Fig. 7. It is evident that  $H$ -probe correctly estimates the configured Hurst parameters.

In a second experiment we inject LRD traffic with differing  $H$  along links 1 and 2 denoted  $H_1$  and  $H_2$  respectively. In Tab. II we show exemplary Hurst parameters obtained for all combinations of  $H_1 = \{0.6, 0.9\}$  and  $H_2 = \{0.6, 0.9\}$ . We note that our method correctly characterizes the dominant correlations, respectively,  $H$  along end-to-end paths from a probing rate of as low as 70 kbps.

#### E. Internet measurements

We perform measurements over multiple weeks using  $H$ -probe from our lab that is connected to the German research network targeting a number of worldwide PlanetLab nodes, in

order to estimate the correlations on end-to-end paths across the Internet. *Traceroute* results show stable paths to each target, e.g., with {15, 17, 16} hops for the targets in Fig. 8(a), 8(b), 8(c) respectively. We provide extended results in the technical report [25]. The complex correlation structure along exemplary Internet paths is illustrated by the covariance plots in Fig. 8. First, we observe LRD covariance decay depicted in Fig. 8(a) and 8(b). We point out that the correlation and hence the Hurst parameter vary significantly throughout the day. Moreover, we find that the correlation structure varies strongly across different paths. Additionally, for some targets we observed distinct periodicities on different timescales, as exemplified in Fig. 8(c). While periodic behavior in offline Internet traces, due to various protocol implementations has been previously reported, e.g., in [5] *H-probe* provides a new tool enabling researchers to shed light on the complex structure of traffic correlations without requiring the availability of traffic traces from Internet service providers.

## V. CONCLUSIONS

In this paper, we derived estimators for the correlations of network traffic, given limited traffic samples obtained by passive monitoring or active probing. We explored the impact of different sampling strategies on observed traffic correlations and quantified the impact of sampling on the observations. We showed that for finite sample sizes there are intrinsic limitations on the accuracy of the estimates and showed the influence of different sampling parameters. We found a non-linear tradeoff between sampling duration and sampling intensity. Further, we inferred the Hurst parameter  $H$  from covariance estimates to quantify LRD. We developed and deployed an active probing method that estimates traffic correlations from end-to-end measurements without network support. The corresponding software is made publicly available. Finally, we presented measurement results from a controlled testbed environment as well as Internet paths. We observe a complex correlation structure on Internet paths. The correlation structure as well as  $H$  significantly vary across time and paths. In addition to LRD we observe periodic behavior at different time scales.

## REFERENCES

- [1] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot. On optimal probing for delay and loss measurement. In *Proc. of IMC*, pages 291–302, 2007.
- [2] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot. The role of PASTA in network measurement. *IEEE/ACM Trans. Netw.*, 17(4):1340–1353, 2009.
- [3] J. Beran. *Statistics for Long-Memory Processes*. Chapman & Hall/CRC, Oct. 1994.
- [4] Z. Bozakov, A. Rizk, and M. Fidler. *H-probe*, 2012. Available at: <http://www.ikt.uni-hannover.de/h-probe>.
- [5] A. Broido, R. King, E. Nemeth, and K. Claffy. Radon spectroscopy of inter-packet delay. In *Proc. of High Speed Networking Workshop*, 2003.
- [6] D. Cox and P. Lewis. *The statistical analysis of series of events*. Methuen's Statistical Monographs, 1966.
- [7] M. Crovella and A. Bestavros. Self-similarity in World Wide Web traffic: evidence and possible causes. *IEEE/ACM Trans. Netw.*, 5(6):835–846, Dec. 1997.
- [8] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. Netw.*, 4(2):209–223, 1996.
- [9] A. Ganesh, N. O'Connell, and D. Wischik. *Big Queues*. Springer, 2004.
- [10] H. Gupta, A. Mahanti, and V. Ribeiro. Revisiting coexistence of poissonson and self-similarity in Internet traffic. In *Proc. of MASCOTS*, pages 1–10, Sept. 2009.
- [11] G. He and J. Hou. On exploiting long range dependency of network traffic in measuring cross traffic on an end-to-end basis. In *Proc. of INFOCOM*, pages 1858–1868, 2003.
- [12] M. Jain and C. Dovrolis. End-to-end available bandwidth: measurement methodology, dynamics, and relation with TCP throughput. *IEEE/ACM Trans. Netw.*, 11(4):537–549, Aug. 2003.
- [13] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of Ethernet traffic. *IEEE/ACM Trans. Netw.*, 2(1):1–15, Feb. 1994.
- [14] P. Loiseau, P. Goncalves, G. Dewaele, P. Borgnat, P. Abry, and P. Primet. Investigating self-similarity and heavy-tailed distributions on a large-scale experimental facility. *IEEE/ACM Trans. Netw.*, 18(4):1261–1274, Aug. 2010.
- [15] S. Machiraju, D. Veitch, F. Baccelli, and J. Bolot. Adding definition to active probing. *Computer Communication Review*, 37(2):17–28, 2007.
- [16] M. Mandjes. *Large Deviations for Gaussian Queues*. Wiley & Sons, 2007.
- [17] B. Melamed and W. Whitt. On arrivals that see time averages. *Oper. Res.*, 38(1):156–172, Feb. 1990.
- [18] I. Norros. On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE J. Sel. Areas Commun.*, 13(6):953–962, Aug. 1995.
- [19] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis. RFC2330 - Framework for IP Performance Metrics. <http://www.rfc-editor.org/rfc/rfc2330.txt>, 1998.
- [20] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *IEEE/ACM Trans. Netw.*, 3(3):226–244, 1995.
- [21] A. Philippe and M. Viano. Random sampling of long-memory stationary processes. *Journal of Statistical Planning and Inference*, 140(5):1110–1124, 2010.
- [22] V. Ribeiro, M. Coates, R. Riedi, S. Sarvotham, B. Hendricks, and R. Baraniuk. Multifractal cross-traffic estimation. In *Proc. of ITC Conference on IP Traffic, Modeling and Management*, Sep. 2000.
- [23] V. Ribeiro, R. H. Riedi, and R. G. Baraniuk. Optimal sampling strategies for multiscale stochastic processes. *IMS Lecture Notes - Monograph Series*, 49, Jan. 2006.
- [24] V. J. Ribeiro, R. H. Riedi, and R. G. Baraniuk. Multiscale queueing analysis. *IEEE/ACM Trans. Netw.*, 14(5):1005–1018, 2006.
- [25] A. Rizk, Z. Bozakov, and M. Fidler. *H-probe: Estimating traffic correlations from sampling and active network probing*. Technical Report arXiv:1208.2870, July 2012.
- [26] M. Roughan. Fundamental bounds on the accuracy of network performance measurements. In *Proc. of SIGMETRICS*, pages 253–264, 2005.
- [27] M. Roughan. A comparison of Poisson and uniform sampling for active measurements. *IEEE J. Sel. Areas Commun.*, 24(12):2299–2312, 2006.
- [28] M. Roughan, D. Veitch, and P. Abry. Real-time estimation of the parameters of long-range dependence. *IEEE/ACM Trans. Netw.*, 8(4):467–478, Aug. 2000.
- [29] J. Strauss, D. Katabi, and F. Kaashoek. A measurement study of available bandwidth estimation tools. In *Proc. of IMC*, pages 39–44, 2003.
- [30] M. Taqqu, V. Teverovsky, and W. Willinger. Estimators for long-range dependence: An empirical study. *Fractals*, 3(4):785–798, 1995.
- [31] M. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *Comput. Commun. Rev.*, 27(2):5–23, Apr. 1997.
- [32] M. B. Tariq, A. Dhamdhere, C. Dovrolis, and M. Ammar. Poisson versus periodic path probing (or, does PASTA matter). In *Proc. of IMC*, pages 119–124, 2005.
- [33] D. Veitch and P. Abry. A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Trans. Inf. Theory*, 45(2):878–897, Apr. 1999.
- [34] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Trans. Netw.*, 5(1):71–86, Feb. 1997.
- [35] R. Wolff. Poisson arrivals see time averages. *Operations Research*, 30(2):223–231, 1981.