# A STOCHASTIC ANALYSIS OF END-TO-END AVAILABLE BANDWIDTH ESTIMATION

by

## XILIANG LIU

A dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy The City University of New York

2005

# Abstract

A STOCHASTIC ANALYSIS OF END-TO-END AVAILABLE

BANDWIDTH ESTIMATION

by

Xiliang Liu

Advisor: Professor K. Ravindran

This thesis presents a theoretical foundation for packet-train available bandwidth estimation in its most general settings. In the first half of our work, we analyze the asymptotic behavior of packet-train probing in a single-hop network path carrying bursty cross-traffic. We examine the asymptotic average of the packet-train output dispersions and its relationship to the input dispersion. We call this relationship the *response curve* of the network path. We show that the real response curve is provably different from that obtained under fluid cross-traffic models in prior work. This difference, which we refer to as *response deviation*, is one of the previously unknown factors that can cause measurement bias in available bandwidth estimation. We show both analytically and experimentally that the response deviation and its consequent measurement bias vanish as the packet-train length or probing packet size increases and that the vanishing rate is decided by the burstiness of cross-traffic.

In the second part of this thesis, we analyze the asymptotic behavior of packet-train probing over a multi-hop network path $\mathcal{P}$ carrying arbitrarily routed bursty cross-traffic flows. We show that the response curve $\mathcal{Z}$ is *tightly*

lower-bounded by its *multi-hop fluid counterpart* $\mathcal{F}$, obtained when every cross-traffic flow on $\mathcal{P}$ is hypothetically replaced with a constant-rate fluid traffic flow of the same average intensity and routing pattern. The real curve $\mathcal{Z}$ asymptotically approaches its fluid counterpart $\mathcal{F}$ as probing packet size or packet train length increases. As an implication of these findings, we show that bursty cross-traffic in multi-hop paths causes negative bias to most existing techniques. This bias can be mitigated using long packet-trains. However, the bias is not completely removable for the techniques that use the portion of the single-hop fluid model that differs from $\mathcal{F}$.

Through our probing analysis, we have achieved a clear understanding of both the validity and the inadequacy of current techniques, and provided a guideline for their further improvements.

# Dedication

To Puisan Wong

# Acknowledgements

# Contents

# List of Figures

# List of Tables

# List of Appendices

# Chapter 1

# Introduction

The information about the spare capacity of an Internet path can be very useful in many applications. With such a knowledge, a TCP connection can quickly ramp up to an optimal sending rate without using slow-start, improving end-to-end throughput; end users can select among several mirror servers for fast downloading; Overlay networks can optimize their application-level topology and provide better services; Network managers can troubleshoot networking problems more efficiently.

However, the available bandwidth information is usually not directly accessible from the network due to the lack of administrative privileges. Therefore, researchers have been developing measurement techniques that infer the bandwidth information from the end points of network paths. This effort can be traced back to early 1990's, and has formed one of the important Internet measurement research area since then.

Related to network bandwidth, there are two concepts that need to be clarified – bottleneck capacity and available bandwidth. The former refers to the maximum rate (in bits per second) a network path can transmit data; while

the later is the spare capacity within a certain time interval after the network path transmits cross-traffic. Path capacity is a static metric determined by the hop with the minimum link capacity; while path available bandwidth is a dynamic metric that relates to the cross-traffic traversing the path. Due to the burstiness of cross-traffic, path available bandwidth varies over time as well as a wide range of observation intervals. We illustrate the two concepts and their relationship using Fig. 1.1, where the height of each rectangle represents the link capacity and the height of shaded portion represents the amount of capacity used to transmit cross-traffic. The height of unshaded portion in each rectangle represents the link available bandwidth. The path available bandwidth is the minimum link available bandwidth along the path. More formal definitions will be given at a later time.
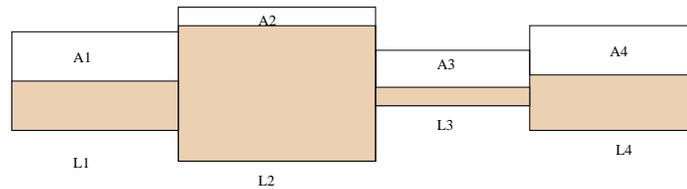


Figure 1.1: Illustration of capacity and available bandwidth.

It is important to notice that the metrics of path available bandwidth usually exhibit a great deal of statistical variability. Consequently, a major issue regarding the practical usefulness of bandwidth estimation is the predictability of this metric based on prior measurements. There were several studies which showed that the available bandwidth along many Internet paths remains stable in relatively long time scales and that measurements conducted back-to-back are expected to produce similar results. Zhang *et al.* measured the TCP throughput of one-megabyte transfers in every minute for a total duration of five hours

[44]. The measurement included 49,000 TCP connections along 145 different Internet paths. They found that the time periods in which the throughput time series can be modeled as a stationary process (and even an $i.i.d$ process ), often last for more than one hour. Balakrishnan $et$ $al.$ examined the throughput stationarity of successive Web transfers to a set of clients [5]. They found that the throughput to a given client appeared to be piece-wise stationary in the timescale of hundreds of minutes.

On the other hand, the amount of time it takes to generate one measurement of the path available bandwidth using current techniques is usually no more than tens of seconds. This measurement result, given that it is sufficiently accurate, can serve as a good prediction for the average available bandwidth of a substantially long time interval (say, twenty minutes) to come. As illustrated in Fig. 1.2, the time interval $[t_0, t_2]$ is the duration when available bandwidth process remains in a stationary state; the interval $[t_0, t_1]$ is the duration to generate one measurement; and $[t_1, t_2]$ is the interval the measurement result remains good. This observation justifies the usefulness of available bandwidth estimation.
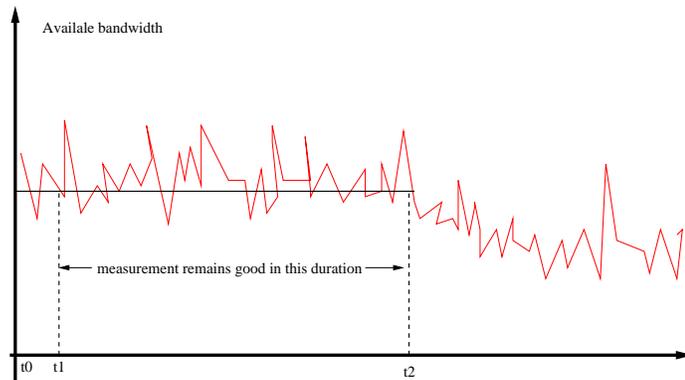


Figure 1.2: Timescales for bandwidth measurements.

End-to-end measurement of bottleneck capacities and available bandwidths involves sending probing packets over the Internet path to infer bandwidth information from the delays or inter-packet delays (i.e., dispersions) of the probing packets when they are received at the destination host. In delay-based measurements, path characteristics such as per-hop capacity and link utilization are inferred based on the RTT or one-way delay of *individual* packets [4], [11], [17], [22], [31]. In dispersion-based measurements, the inter-packet delays of packet-pairs is traditionally used to infer bottleneck capacity [7], [8], [9], [12], [16], [21], [33], [34]; however, recent approaches also use packet-pairs/trains to measure cross-traffic and available bandwidth of an end-to-end path [6], [13], [15], [18], [28], [36], [38]. It is straightforward to understand and theoretically validate delay-based bandwidth measurements. The major difficulties remained are mostly due to practical issues [31], [35]. On the other hand, it is far more difficult to characterize dispersion-based methodology that involves packet pair/train probing. Consequently, apart from the practical issues, dispersion-based measurement techniques are yet to be fully justified for general path and cross-traffic conditions.

There has been a fair amount of research effort to achieve a clear understanding of packet-train bandwidth estimation. However, previous analysis either relied on constant-rate fluid cross-traffic models [9], [27], or provided answers that are restricted to a single-hop path and also are only *partially* suitable for generic bursty cross-traffic. [7], [15], [32], [36]. A packet-train bandwidth measurement theory in the context of a multi-hop path with general bursty cross-traffic arrival still remains as an open problem. Such a theory is important in that it helps to understand both the validity and the inadequacy of existing techniques and provides a guideline for their further improvement. In this thesis, we establish

such a theory for end-to-end packet-train available bandwidth estimation.

Before getting into the detailed statement of our research problems, it is necessary to briefly discuss the differences between available bandwidth measurement and capacity measurement. Both of the two types of measurements are based on packet-train probing, and can be classified as dispersion-based techniques. However, capacity measurement techniques measure the minimum link capacity along the path. They largely rely on various heuristics to detect the mode that corresponds to the minimum capacity from the *distribution* (i.e., histogram) of the output packet-pair dispersions. On the other hand, available bandwidth measurement techniques measure the spare capacity of a network path. They are mostly based on the *statistical mean* of the output packet-train dispersions and its mathematical relationship to the input packet-train dispersion. This thesis focuses on the later, and not the former.

## 1.1  Research Problem

The central problem in packet-train bandwidth estimation is to derive the functional relation between the input and output packet-train dispersions and to show how the path bandwidth information is related to this functional relation. In a multi-hop path with arbitrary cross-traffic, this problem appeared to be very challenging. Therefore, as a natural starting point, previous work used a single-hop path with constant-rate fluid (CRF) cross-traffic to simplify analysis. CRF cross-traffic is a hypothetical traffic with infinitely small packet size and constant arrival rate $\lambda$. For any time interval $[t, t + \delta]$, the amount of CRF traffic received by the network path is $\lambda \times \delta$. The deterministic and constancy nature of CRF cross-traffic leads to a simple closed-from expression of packet-

train output dispersion $g_O$ as a piece-wise linear function of the input dispersion $g_I$, as given in the following

$$g_O = \max\left(g_I, \frac{s + \lambda g_I}{C}\right) = \begin{cases} g_I & g_I \geq \dfrac{s}{C - \lambda} \\ \dfrac{s + \lambda g_I}{C} & g_I \leq \dfrac{s}{C - \lambda} \end{cases}, \qquad (1.1)$$

where $s$ is the probing packet-size and $C$ is the hop capacity. Letting $r_I = s/g_I$ and $r_O = s/g_O$ be the input and output probing rate respectively, we get the rate version of (1.1) as follows

$$r_O = \min\left(r_I, \frac{r_I C}{r_I + \lambda}\right) = \begin{cases} r_I & r_I \leq C - \lambda \\ \dfrac{r_I C}{r_I + \lambda} & r_I \geq C - \lambda \end{cases}. \qquad (1.2)$$

Another commonly used variation of (1.2) is given by the piece-wise linear function between $r_I/r_O$ and $r_I$

$$r_I/r_O = \max\left(1, \frac{r_I + \lambda}{C}\right) = \begin{cases} 1 & r_I \leq C - \lambda \\ \dfrac{r_I + \lambda}{C} & r_I \geq C - \lambda \end{cases}. \qquad (1.3)$$

It is conceptually helpful to view the probing output (dispersion or rate) as the *response* of the network path to the probing input. We call the functional relation between the response and the input the "probing response curve" of the network path. Most existing techniques for available bandwidth estimation are based upon or related to the single-hop fluid response curves (1.1), (1.2), or their variants such as (1.3). In practice, a network path usually consists of multiple hops and cross-traffic are always bursty. Note that due to the random nature of bursty cross-traffic, the output packet-train dispersions are also random. Accordingly, the statistical mean of the output dispersions is viewed as the *response* of the network path to the input dispersion. Current measurement proposals assume that the impact of non-bottleneck links is negligible and

that cross-traffic burstiness only causes measurement variability that can be smoothed out by averaging multiple probing samples. In other words, without formal justification, the single-hop fluid response curve is considered to be a valid approximation of the multi-hop response curve in bursty cross-traffic.

In this thesis, we set our research goals to address the following questions:

- Derive the response curve for an arbitrary network path with bursty cross-traffic arrival.

- Investigate the impact of input packet-train parameters (i.e., probing packet size and packet-train length) on the probing response curves.

- Examine the impact of cross-traffic burstiness on the probing response curves. In a multi-hop path, examine the impact of cross-traffic routing on the probing response curves.

- Compare the real response curve with the one obtained in CRF cross-traffic, and analyze the validity and inadequacy of current techniques.

In summary, this thesis establishes a theory for packet-train bandwidth estimation by deriving the "probing response curves" and uncovering their major properties.

## 1.2   Solution

We tackle the questions listed above in three steps. First, we take a sample-path approach to analyze packet-train bandwidth estimation in a single-hop path with bursty cross-traffic arrival. The results obtained in this step (see chapter 3) are important building blocks for later multi-hop analysis.

Even though many single-hop results can be treated as special cases in multi-hop analysis, there are several reasons that we need to separate them from the multi-hop analysis. First, due to the different levels of complexity, this thesis treats single-hop analysis with a different level of mathematical rigor than it treats multi-hop analysis. We make much weaker assumptions and use fewer approximations in the single-hop analysis. Hence, our results have a boarder applicability in practice. While in the multi-hop analysis, we rely on a stronger assumption on cross-traffic arrival and an additional approximation on cross-traffic departure at each hop to deal with the increased complexity of the problem, as we shall see in chapter 4. The second reason we separate the single-hop analysis is that even in a single-hop case, the problem is fairly complex and a thorough treatment requires a significant amount of effort.

The main results we obtained in the first step is a closed-form expression for single-hop probing response curve. Our results show that cross-traffic burstiness causes the real response curve to deviate from that obtained in CRF cross-traffic. This response deviation may introduce significant measurement bias to existing techniques. We also find that by increasing packet-train length, the response deviation and measurement bias it causes can be reduced to a negligible level.

In the second step, we derive the multi-hop response curve $\mathcal{F}$ for a network path that carries arbitrarily routed CRF cross-traffic flows (see details in the second section of chapter 4). We obtain a recursive closed-form expression of $\mathcal{F}$.

In the third step, we derive the multi-hop response curve $\mathcal{Z}$ for a network path $\mathcal{P}$ that carries arbitrarily routed bursty cross-traffic flows (in the third and fourth sections in chapter 4) and compare it to its "multi-hop fluid counterpart" $\mathcal{F}$, which is obtained when every cross-traffic flow in $\mathcal{P}$ is hypothetically replaced

with a CRF flow of the same arrival rate and routing pattern.

In addition to the response deviation $\mathcal{Z} - \mathcal{F}$ caused by cross-traffic burstiness, we find another source of response deviation that comes from the difference between the multi-hop CRF response $\mathcal{F}$ and the single-hop (the bottleneck link) CRF response $\mathcal{S}$, which the current techniques are anchoring upon. Both types of response deviation are positive and cause current techniques to underestimate the path available bandwidth. The portion of underestimation caused by the response deviation $\mathcal{Z} - \mathcal{F}$ is elastic, meaning that it can be reduced to a negligible level using long packet-trains. On the other hand, the portion of underestimation caused by the response deviation $\mathcal{F} - \mathcal{S}$ is non-elastic and remains constant for arbitrary packet-train parameters. The way to keep away from non-elastic measurement bias is to avoid probing the path at high input rate.

## 1.3   Contributions

This thesis makes the following contributions:

- It provides a stochastic-theoretic characterization of packet-train bandwidth estimation and uncovers the fact that cross-traffic burstiness introduces measurement bias (in addition to measurement variabilities) to current techniques.

- It propose a methodology to compute the single-hop response curve from a given cross-traffic trace.

- It proves that ignoring multi-hop effect in bandwidth estimation can also lead to measurement bias.

- It gives a complete answer to the question regarding the implication of cross-traffic routing on bandwidth estimation, which has been mostly over-looked in prior work.

- It formally proves that the measurement bias caused by cross-traffic bursti-ness can be overcome by increasing packet-train parameters.

- It leads to a new measurement method that estimates the utilization and capacity of the bottleneck link in multi-hop network paths.

These contributions lead to a fairly complete understanding of the packet-train bandwidth estimation problem and the fundamental tradeoffs therein.

## 1.4   Dissertation Overview

This dissertation is organized as follows. In chapter 2, we give a brief survey of current bandwidth estimation techniques and point out that the rationales they are anchoring upon are all related to (1.1). This motivates us to examine the validity of (1.1) in a network path with bursty cross-traffic. In chapter 3, we take a sample-path approach to derive the probing response curve in a single-hop path with bursty cross-traffic arrival. We demonstrate the deviation of the single-hop response curve from (1.1) and the measurement bias it causes to existing techniques. In chapter 4, we give a stochastic analysis on the response curve $\mathcal{Z}$ of a multi-hop path and decompose its deviation from (1.1) into two portions. We demonstrate the measurement biases caused by both portions of response deviation and point out the way to mitigate or avoid the two types of consequent measurement biases. Finally, in chapter 5, we summarize our work and point out future research directions.

# Chapter 2

# Related Work

IP-layer bandwidth estimation and the idea of using packet-pairs to infer link capacity dates at least as far back as 1988 when Jacobson [16] designed the packet conservation principle of TCP to allow senders to indirectly infer the bottleneck/available bandwidth based on the spacing between the ACK packets. Keshav's packet-pair flow control followed in 1991 [20] and relied on fair queuing in all network routers.

Several years later, Carter *et al.* (1996) developed a tool called `cprobe` [8] to measure the available bandwidth. `Cprobe` bounced a short train of ICMP echo packets off the target server and recorded the spacing between the first and last returning packet. The rate of the arriving echo stream was used as an estimate of the available bandwidth. As pointed out later by Dovrolis [9], `cprobe` actually measured a metric called the *asymptotic dispersion rate* (ADR), which does *not* generally equal the available bandwidth. Paxon (1999) defined and measured a relative available bandwidth metric $\beta$ [34], which approached 1 when the path was void of cross-traffic and 0 when the path was close to 100% utilization.

Melander *et al.* (2002) studied the relationship between the input and output

rates $r_I$ and $r_O$ of probing trains in a single-hop path and presented the following FIFO fluid model [27]:

$$r_O = \begin{cases} r_I & r_I \leq C - \lambda \\ C\dfrac{r_I}{r_I + \lambda} & r_I \geq C - \lambda \end{cases}, \tag{2.1}$$

where $C$ and $\lambda$ are the hop capacity and cross-traffic intensity (or rate) respectively. Applying math induction to the subsequent hops along the path, we get the main model of measuring the available bandwidth $A_P$ of an arbitrary multi-hop path $P$:

$$r_O = \begin{cases} r_I & r_I \leq A_P \\ C\dfrac{r_I}{r_I + \lambda} & b \geq r_I \geq A_P \end{cases}, \tag{2.2}$$

where $b$ is the second minimum residual link bandwidth along path $P$ and $C$ is the capacity of the tight hop.

Based on (2.1) and (2.2), Melander *et al.* proposed a measurement technique called TOPP (Trains of Packet Pairs) [28]. TOPP first collects the output rates of probing packet pairs for a series of equally spaced input rates in some interval $[r_I^{min}, r_I^{max}]$. In the subsequent analysis phase, instead of using (2.2), TOPP uses the piece-wise linear relationship between $r_I/r_O$ and $r_I$:

$$\frac{r_I}{r_O} = \begin{cases} 1 & r_I \leq A_P \\ \dfrac{r_I}{C} + \dfrac{\lambda}{C} & b \geq r_I \geq A_P \end{cases}. \tag{2.3}$$

TOPP identifies the second segment in the curve using several empirical methods and applies linear regression to calculate the capacity $C$ and cross traffic intensity $\lambda$ of the tight link. Hence, $A_P = C - \lambda$ is obtained.

Another recent proposal is SLoPS (Self Loading Periodic Streams) by Jain *et al.* (2002) [18]. SLoPS is implemented in a tool called `pathload` and is based

on the observation that one-way delays of packets in a probing train show an increasing trend when the input rate of the probe traffic is higher than the available bandwidth of the path. This rationale is clearly true if cross-traffic is modeled as a fluid and generally can be written as a variation of (2.3):

$$\frac{r_I}{r_O} = \begin{cases} 1 & r_I \leq A_P \\ > 1 & r_I > A_P \end{cases}. \tag{2.4}$$

To measure available bandwidth in bursty cross-traffic, `pathload` adapts its input probing rate in a way similar to a binary search to locate the region where the one-way delay of the probing packets is just about to show an increasing trend or the two statistical tests used can neither detect an increasing trend, nor detect a non-increasing trend with sufficient confidence. That region is then taken as the range of the available bandwidth of the path.

`PathChirp` [37] is a proposal to improve `pathload`'s measurement speed. `PathChirp` uses probing trains with exponentially decreasing inter-packet spacing and calculates available bandwidth from the *queuing delay signature* of the arriving chirp.

Hu *et al.* [15] (2003) analyzed the interaction between probing pairs and CBR cross-traffic using a single-hop path. They proposed the following gap formula under the condition that the packets in each probing pair share the same hop busy period:

$$g_O = \frac{s}{C} + \frac{\lambda g_I}{C}, \tag{2.5}$$

where $g_O$ is the output gap, $g_I$ is the input gap between the packet pair, $s$ is the packet size of probe traffic. The paper [15] also proposed a packet-train based estimator called IGI that measures the cross-traffic intensity, which can be viewed as an empirical extension of (2.5).

As an alternative to IGI, [15] suggested to use a method called PTR (Packet Transmission Rate), in which the output rate of the probing train is used as an estimator of $A_P$. The authors [15] showed that both IGI and PTR produce accurate results at the *turning point* where the input gap $g_I$ *starts* to become the same as the output gap $g_O$.

Notice that IGI/PTR is also related to model (2.1), which shows that the *turning point* is where both $r_I$ and $r_O$ are equal to the available bandwidth $C - \lambda$. Equation (2.5) is the "gap" version of the second part of (2.1).

Spruce [38] is another measurement proposal that uses packet-pairs. Like IGI, spruce assumes a single bottleneck link whose capacity $C$ can be estimated beforehand. Spruce sends probing pairs with intra-pair gap $g_I$ set to the bottleneck link transmission delay of the packet and inter-pair delay set to an exponentially distributed random variable so as to maintain the average probing rate below $0.05C$. Each probing pair generates an available bandwidth estimate $A_i$ computed by:

$$A_i = C\Big(1 - \frac{g_O - g_I}{g_I}\Big). \tag{2.6}$$

Spruce averages the last 100 samples of $A_i$ to arrive at an estimation of $A_P$. Observe that spruce anchors its rationale on (2.5) with $g_I = s/C$, where $s$ is the probing packet size.

There are other measurement proposals such as Delphi [36] and the work in [13]. These proposals assume specific cross-traffic processes, which allows them to either directly estimate cross-traffic intensity or reconstruct its parameters on a larger timescale based on the sampled traffic in small time intervals. The packet probing part however is similar to that of spruce and is related to (2.5).

In summary, most of the recent proposals anchor their rationales directly on

(2.1) or a model closely related to it. However, (2.1) is only fully justified based on a fluid cross-traffic model, in which the arrival rate of cross-traffic is constant at all times $t$ and equals $\lambda$. For general bursty cross-traffic, it is important to understand whether (2.1) is the asymptotic behavior of packet train probing or not. An affirmative answer to this question would lay a solid ground for the design of available bandwidth measurement methods and provide them with an assurance of asymptotic accuracy. On the other hand, a negative answer would shed new light on the fundamental limits and tradeoffs in probing-based measurements, giving rise to new insights in parameter tuning under certain application requirements. In the next chapter, we tackle this question in a single-hop path.

# Chapter 3

# Single-Hop Probing Analysis

## 3.1   Introduction

Let us start from a more accurate definition of available bandwidth. According
to recently established notions, the available bandwidth of a network hop is its
*residual* capacity after transmitting cross-traffic. Since at any time instance, the
hop is either idle or transmitting packets at its capacity speed $C$, the utilization
of the hop can be viewed as an on-off function over time. The definition of
the available bandwidth ought to look at the average unutilized bandwidth over
some time interval $\delta$, i.e.,

$$B_\delta(t) = C\Big(1 - \frac{1}{\delta} \int_t^{t+\delta} U(x)dx\Big), \qquad (3.1)$$

where $B_\delta(t)$ is the available bandwidth in time interval $[t, t+\delta]$, $U(x) \in \{0, 1\}$
is the link utilization on-off function determined by the packet-arrival pattern
of cross-traffic, and $C$ is the hop capacity. The available bandwidth along a
network path is the minimum available bandwidth of all traversed hops. The
hop carrying the minimum available bandwidth is called the *tight hop*.

In this chapter, we analyze the asymptotic behavior of single-hop, packet-train bandwidth estimation under bursty cross-traffic conditions. This question has two aspects. First, given a cross-traffic arrival process and fixed probing train parameters (i.e., packet size and train length), we analyze how the probing output relates to the probing input. We investigate the output rate and gap for individual packet trains as well as their asymptotic average as the number of probings approaches infinity. We examine the functional relation between the probing input and the asymptotic average of the probing output in the entire input range.

Second, we investigate how the response curve evolves with respect to the changes in packet train parameters and cross-traffic burstiness. Both questions are of fundamental importance for the design of available-bandwidth estimation methods. The answer to the first question provides a theoretical foundation that extends previous rationales based on fluid cross-traffic models. The answer to the second question offers an insight into parameter tuning strategies in the measurement design.

Although our final goal is to understand the behavior of packet-train probing in multi-hop network paths, the insight obtained in the analysis of a single hop is indispensable in reaching this goal. Moreover, the single-hop case on its own is an interesting and complex problem calling for an elaborate discussion.

In this chapter, we make two theoretically and practically mild assumptions, under which we derive several important properties of the gap (and rate) response curve. Our results show that the rate response curve in constant-rate fluid cross-traffic is the tight upper bound of that in bursty cross-traffic with the same average intensity. We show that there is a probing input range where the real curve negatively deviates from its fluid-based prediction. Most exist-

ing measurement techniques make use of the curve in that range without being aware of the "response deviation", which sometimes makes them subject to significant measurement bias.

Our analysis also identifies the source of the probing response deviation and arrives at its closed-form expression for arbitrary packet-train parameters. We show that the amplitude of the response deviation is exclusively decided by the packet-train parameters and the available bandwidth distribution. We also present an experimental approach to compute the response deviation in given cross-traffic traces. This allows us to empirically validate our theoretical results, qualitatively observe the relationship between the response deviation and probing train parameters in certain cross-traffic conditions, and evaluate the asymptotic performance of various available-bandwidth estimators.

The rest of the chapter is organized as follows. In section 3.2, we identify the measurement targets and present the analytical foundation of packet-train probing. In section 3.3, we analyze the major properties of the response curves. In section 3.4, we propose two experimental methods, period testing and trace driven testing, to observe the response deviation and examine its relationship to several deciding factors. We explain the implications of our findings on some of the current proposals in section 3.5. Finally, we present the concluding remarks in section 3.6.

## 3.2 Analysis of Packet Probing

In this section, we present an analytical formulation of packet probing, identify measurement targets, and derive closed-form relation between probing input and output for individual packet trains. Our analysis focuses on the single-hop

input cross-
traffic

output cross-
traffic

input probe
traffic

router of capacity $C$

output probe
traffic
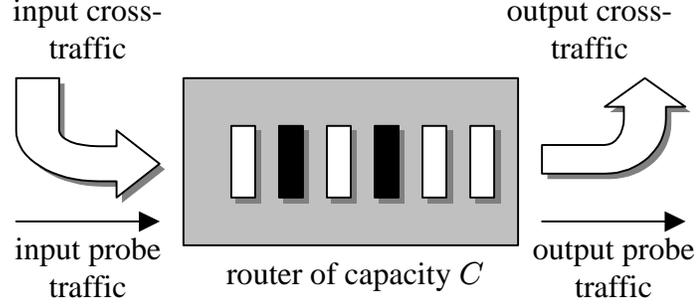
Figure 3.1: Single-hop probing model.

probing model in Figure 3.1. We use the quadruple $\langle a_1, g_I, s, n \rangle$ to denote a probing train of $n$ packets $p_1, p_2, \ldots, p_n$, where $a_1$ is the arrival time of the first packet $p_1$ to the hop, $g_I$ is the inter-packet spacing, $s$ is the probe packet size, and $n$ is the train length. The arrival time at the hop of the probing packets are denoted by $a_i = a_1 + (i-1)g_I, i = 1, 2, \ldots, n$. The departure time of probing packets from the hop are denoted by $d_i, i = 1, 2, \ldots, n$. We define the *output gap* of a packet train as the *average* spacing between adjacent packets in the train :

$$g_O = \frac{d_n - d_1}{n - 1}. \tag{3.2}$$

In terms of rate, the corresponding average *input* and *output* rates are given by:

$$r_I = \frac{s}{g_I}, \qquad r_O = \frac{s}{g_O} = \frac{(n-1)s}{d_n - d_1}. \tag{3.3}$$

We start from the gap version of (2.1), namely, we first investigate the validity of the following model:

$$E[g_O] = \begin{cases} g_I & g_I > \dfrac{s}{C - \lambda} \\ \dfrac{s}{C} + \dfrac{g_I \lambda}{C} & g_I \leq \dfrac{s}{C - \lambda} \end{cases} \tag{3.4}$$

in a single hop path and then come back to its rate version. Since we are now dealing with bursty cross-traffic, neither cross-traffic intensity nor probing output gap is a constant. Meanwhile, $\lambda$ and $E[g_O]$ can be viewed as the time averages of traffic intensity and output gaps. Detailed connotations about these two terms are clarified at later proper times.

### 3.2.1  Problem Formulation

Throughout this chapter, we assume infinite buffer capacity, FIFO queuing, and a work-conserving discipline for the forwarding hop. For the composition of cross-traffic and probing traffic, we assume simple traffic arrival, i.e., at most one packet arrives at any time instance.

**Definition 1** *Cross traffic is driven by the packet counting process $N(t)$ and the packet-size process $S_n$. The cumulative traffic arrival $V(t)$ is a random process counting the total volume of data received by the router up to time instance $t$:*

$$V(t) = \sum_{n=1}^{N(t)} S_n. \tag{3.5}$$

Note that $V(t)$ and $N(t)$ are right continuous, meaning that the packet arriving at $t$ is counted in $V(t)$. Unlike conventional traffic modeling, we make no assumption on $N(t)$ or $S_n$. Instead, our assumption is made for $V(t)$.

**Assumption 1** *Cross traffic exhibits "intensity stability," which means that $\lim_{t\to\infty} V(t)/t$ exists and is less than the hop capacity $C$.*

This higher level assumption can accommodate a broad range of traffic types and, at the same time, detach the model from the underlying details of traffic arrival. We define cross-traffic intensity $\lambda$ in (3.4) as the limit of $V(t)/t$ as

$t \to \infty$. This definition reveals a mathematical essence of one's intuitive notion of average traffic intensity. Further, as we next show, the time average of cross-traffic intensity metrics in *arbitrary* fixed observation interval is the same as this limit.

**Definition 2** *We define $Y_\delta(t)$ as the average cross-traffic arrival rate in the interval $(t, t + \delta]$ and call it the "$\delta$-interval cross-traffic intensity" process:*

$$Y_\delta(t) = \frac{V(t + \delta) - V(t)}{\delta}. \tag{3.6}$$

Given this definition, we have the following result.

**Lemma 1** *The limiting time average $E[Y_\delta(t)]$ of any $\delta$-interval cross-traffic intensity sample-path is equal to $\lambda$:*

$$E[Y_\delta(t)] = \lim_{t \to \infty} \frac{1}{t} \int_0^t Y_\delta(u)du = \lambda, \quad \forall \delta > 0. \tag{3.7}$$

**Proof:** First, notice that:

$$\frac{1}{t} \int_0^t Y_\delta(u)du = \frac{\int_t^{t+\delta} V(u)du}{\delta t} - \frac{\int_0^\delta V(u)du}{\delta t}. \tag{3.8}$$

Computing the limits, we get:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t Y_\delta(u)du = \lim_{t \to \infty} \frac{\int_t^{t+\delta} V(u)du}{\delta t} - 0. \tag{3.9}$$

Since $V(t)$ is a non-decreasing function, we can write:

$$\delta V(t) \le \int_t^{t+\delta} V(u)du \le \delta V(t + \delta). \tag{3.10}$$

Finally, note that both $\delta V(t)$ and $\delta V(t+\delta)$ have the same limit when divided by $\delta t$:

$$\begin{aligned}
\lim_{t \to \infty} \frac{V(t)}{t} &= \lim_{t \to \infty} \frac{\delta V(t)}{\delta t} \le \lim_{t \to \infty} \frac{\int_t^{t+\delta} V(u)du}{\delta t} \\
&\le \lim_{t \to \infty} \frac{\delta V(t + \delta)}{\delta t} = \lim_{t \to \infty} \frac{V(t + \delta)}{t + \delta} \frac{t + \delta}{t} \\
&= \lim_{t \to \infty} \frac{V(t)}{t}.
\end{aligned} \tag{3.11}$$

Combining (3.9) and (3.11), we have for $\forall \delta > 0$:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t Y_\delta(u)du = \lim_{t \to \infty} \frac{V(t)}{t} = \lambda, \qquad (3.12)$$

which leads to the statement of the lemma. ∎

Throughout this chapter, we use the notation of probability expectation as a shorthand representation for sample-path limiting time average. In fact, the limiting time average of a sample-path is the expectation of its limiting frequency distribution [29, pages 45-50]. Hence, it is also called the "sample-path mean." The analysis in this chapter is purely sample-path based, and we avoid addressing any probabilistic nature of the underlying random process. The first equality in Lemma 1 has nothing to do with ergodicity. It is an equality by definition. Lemma 1 reveals that to measure $\lambda$, instead of conducting one observation in a very large time interval (which is often not practical), we can conduct observations in arbitrarily small time intervals and use their average to approach it. This has significant implication on probing based measurement as we show later.

Our next assumption is related to the forwarding hop.

**Definition 3** *Hop workload process $W(t)$ is the sum at time instance $t$ of service times of all packets in the queue and the remaining service time of the packet in service.*

Note that $W(t)$ is also right continuous. Two examples of hop workload sample-path are shown in Figure 3.2.

**Assumption 2** *The forwarding hop exhibits workload stability. That is, $\lim_{t \to \infty} W(t)/t = 0$.*

(a) Exponential on-off          (b) CBR
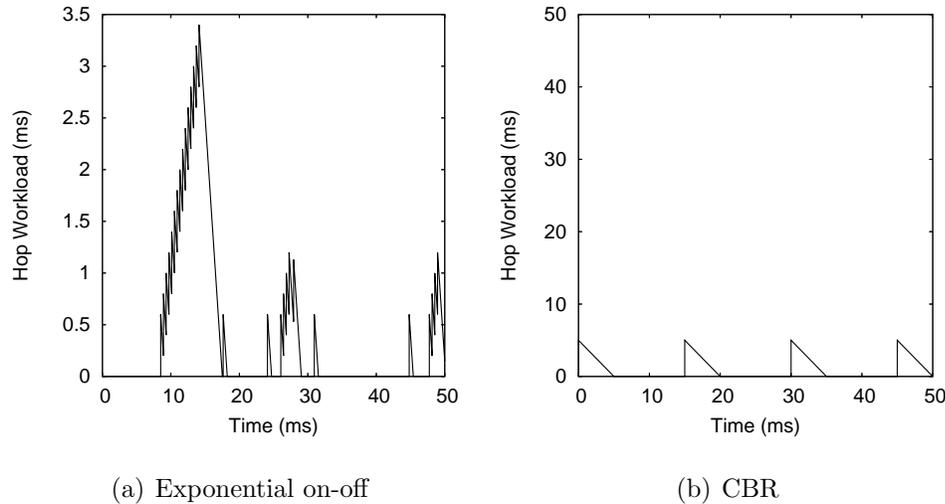
Figure 3.2: (a) The first 50 ms of the workload sample-path $W(t)$ of exponential on-off ns-2 traffic ($C = 10$ mb/s, $s = 750$ bytes); (b) Hop workload $W(t)$ of CBR ns-2 traffic ($C = 2.4$ mb/s, $s = 1500$ bytes).

Workload stability means that $W(t) = o(t)$. Note that given Assumption 1, workload stability is satisfied in most practical situations and that Assumption 2 is formally stated only for convenience of presentation.

We next define a process especially useful in characterizing how cross-traffic changes the gaps of probing packet pairs.

**Definition 4** *A $\delta$-interval workload-difference process $D_\delta(t)$ is the difference between the hop workload at time $t$ and $t + \delta$:*

$$D_\delta(t) = W(t + \delta) - W(t). \tag{3.13}$$

One important implication of workload stability relevant to probing based measurements is the zero-mean nature of $D_\delta(t)$. It is formally stated as follows.

**Lemma 2** *Assuming $W(t) = o(t)$, the limiting time average $E[D_\delta(t)]$ of any $\delta$-interval workload-difference sample-path is zero:*

$$E[D_\delta(t)] = \lim_{t \to \infty} \frac{1}{t} \int_0^t D_\delta(u) du = 0, \quad \forall \delta > 0. \tag{3.14}$$

**Proof:** By the definition of $D_\delta(t)$, we have

$$
\begin{aligned}
E[D_\delta(t)] &= E[W(t+\delta) - W(t)] \\
&= E[W(t+\delta)] - E[W(t)] \\
&= \lim_{t \to \infty} \frac{\int_0^t W(u+\delta) du}{t} - \lim_{t \to \infty} \frac{\int_0^t W(u) du}{t} \\
&= \lim_{t \to \infty} \frac{\int_t^{t+\delta} W(u) du}{t} - \lim_{t \to \infty} \frac{\int_0^\delta W(u) du}{t} \\
&= \lim_{t \to \infty} \frac{\int_t^{t+\delta} W(u) du}{t} - 0 = 0.
\end{aligned}
\tag{3.15}
$$

The last equality holds since $W(t) = o(t)$. ∎

With these two assumptions, we next present a formulation of "available bandwidth" and show how it is related to cross traffic and hop workload.

**Definition 5** *Hop utilization process $U(t)$ is an on-off process associated with $W(t)$:*

$$U(t) = \begin{cases} 1 & W(t) > 0 \\ 0 & W(t) = 0 \end{cases} \tag{3.16}$$

*and $\delta$-interval hop idle process*

$$I(t, t+\delta) = I_\delta(t) = \delta - \int_t^{t+\delta} U(x) dx \tag{3.17}$$

*is a process indicating the total amount of idle time of the forwarding hop in $[t, t+\delta]$. We further call time interval $[t, t+\delta]$ a "hop busy period" if $I_\delta(t) = 0$ and a "hop idle period" if $I_\delta(t) = \delta$.*

Under this picture, several properties of the workload sample-path $W(t)$ for non-fluid traffic become clear. First, $W(t)$ consists of alternating idle and busy periods. Second, in any busy period, $W(t)$ is a series of piecewise linear segments with slope $-1$ separated by discontinuity points. Third, any discontinuous point $d$ in $W(t)$ corresponds to the arrival of a packet. Assuming the packet size is $s$, we have[1] $W(d) - W^-(d) = s/C$.

**Definition 6** *A $\delta$-interval available bandwidth process $B_\delta(t)$ is a process indicating the residual bandwidth in the time interval $[t, t+\delta]$:*

$$B_\delta(t) = C\left(1 - \frac{1}{\delta}\int_t^{t+\delta} U(x)dx\right) = \frac{I_\delta(t)C}{\delta}. \tag{3.18}$$

In our next lemma, we present the relationship among cross-traffic intensity, hop workload, and available bandwidth in arbitrary finite time intervals.

**Lemma 3** *For all $t \geq 0$ and $\delta > 0$, the following holds:*

$$\delta = \frac{B_\delta(t)\delta}{C} - D_\delta(t) + \frac{Y_\delta(t)\delta}{C}. \tag{3.19}$$

**Proof:** Note that the total hop idle time in $[t, t+\delta]$ is

$$I_\delta(t) = \frac{B_\delta(t)\delta}{C}. \tag{3.20}$$

The amount of data transmitted by the hop in $[t, t+\delta]$ is given by the workload change in the hop (taking into account the new arrivals):

$$(W(t) - W(t+\delta))C + V(t+\delta) - V(t)$$
$$= -D_\delta(t)C + Y_\delta(t)\delta, \tag{3.21}$$

---

[1] $f^-(a)$ denotes the left-sided limit $\lim_{x \to a-} f(x)$.

which follows from the definitions of $D_\delta$ and $Y_\delta$ in (3.13) and (3.6). Dividing (3.21) by $C$, the hop working time is

$$-D_\delta(t) + \frac{Y_\delta(t)\delta}{C}. \qquad (3.22)$$

Since the sum of hop working time in (3.22) and hop idle time in (3.20) must be equal to $\delta$, we immediately get the statement of the lemma. ∎

Note that the term $D_\delta(t)$ escaped the formulation efforts of prior work. Although it is a zero-mean term, it is not always insignificant. For example, when the *distribution* of available bandwidth is of interest, this term must be taken into consideration.

The next two theorems present the asymptotic relationship between cross-traffic intensity and available bandwidth. They explain when and why available bandwidth can be estimated by measuring cross-traffic intensity $\lambda$.

**Theorem 1** *Under the assumptions made in this chapter, $\delta$-interval available bandwidth converges to $C - \lambda$ as the observation interval becomes large:*

$$\lim_{\delta \to \infty} B_\delta(t) = C - \lambda, \quad \forall t > 0. \qquad (3.23)$$

**Proof:** Rearranging (3.19), we get:

$$B_\delta(t) = C - Y_\delta(t) + \frac{D_\delta(t)C}{\delta}. \qquad (3.24)$$

Note that since we assumed $W(t) = o(t)$, we have:

$$\lim_{\delta \to \infty} \frac{D_\delta(t)}{\delta} = \lim_{\delta \to \infty} \left( \frac{W(t + \delta) - W(t)}{\delta} \right) = 0. \qquad (3.25)$$

Further, as an immediate consequence of Assumption 1, we have:

$$\lim_{\delta \to \infty} Y_\delta(t) = \lambda, \quad \forall t. \qquad (3.26)$$

Taking the limit of (3.24) and combining with (3.25) and (3.26), we get (3.23).

∎

Theorem 1 shows that given the two stability assumptions we made, available bandwidth also exhibits stability and, in *large* time intervals, can be approximated by $C - \lambda$.

Note, however, that in cases when we are interested in the available bandwidth in a *small* $\delta$-interval[2], Lemma 3 suggests that $B_\delta(t)$ *cannot* be correctly estimated based on the measurement of $Y_\delta(t)$ alone. However, the following theorem says that the limiting time average of available bandwidth metrics in arbitrary $\delta$-interval can be estimated by measuring cross-traffic.

**Theorem 2** *The limiting time average $E[B_\delta(t)]$ of any $\delta$-interval available bandwidth process is $C - \lambda$. That is,*

$$E[B_\delta(t)] = \lim_{t \to \infty} \frac{1}{t} \int_0^t B_\delta(u)du = C - \lambda, \quad \forall \delta > 0. \tag{3.27}$$

**Proof:** This is a direct consequence from (3.24), Lemma 1 and Lemma 2. We leave the verification to the reader. ∎

To summarize, our results show that *available bandwidth in a large timescale or the first-order statistics of available bandwidth in arbitrary fixed time scale can be estimated based on the measurement of cross traffic, while small timescale metrics and their higher-order statistics cannot be correctly estimated solely based cross-traffic measurements.*

Note that measuring cross-traffic intensity $\lambda$ is not the only way to estimate available bandwidth $A$. Metric $A = C - \lambda$ can be directly estimated without knowing the values of $C$ or $\lambda$, as is the case of SLoPS [18] and PTR [15].

---

[2] "Small" is relative to the convergence delay of $V(t)/t$.

Our discussion of probing response curve in Section 4 will cover the theoretical aspects of both approaches.

Despite the perplexing dynamics, we identified two measurement targets, $\lambda$ and $A = C - \lambda$, under mild traffic assumptions. These two targets are fairly stable in the sense that they are independent of any particular observation time instance $t$ and observation interval $\delta$. Although other metrics such as the variance and distribution of available bandwidth might also be interesting, they are less stable because of their dependence on $\delta$. Measurement of those targets is beyond the scope of this thesis.

We are now ready to derive the probing response curve and show how these two targets, $\lambda$ and $A$, are captured in the curve. Before that, however, we must understand the interaction between the probing traffic and the cross-traffic. Traffic interaction includes two parts: the way the probing train changes the original hop workload and the way the cross-traffic changes the inter-packet gaps in the probing train. The latter is our interest, but its analysis relies on understanding the former.

### 3.2.2   Probing Intrusion of Packet Trains

We use $\tilde{W}(t)$ and $\tilde{I}(t)$ to respectively denote the workload sample-path and the hop idle sample-path associated with the superposition of cross-traffic and probing traffic. Note that traffic composition only increases hop workload. That is, for all $t$, $\tilde{W}(t) \geq W(t)$. We next define useful notation that will help us examine this intrusion behavior of packet train probing.

**Definition 7** *The* intrusive range *of the probing traffic into* $W(t)$*, is the set* $\{t : \tilde{W}(t) > W(t)\}$*. The* intrusion residual function *is* $W_d(t) = \tilde{W}(t) - W(t)$*.*
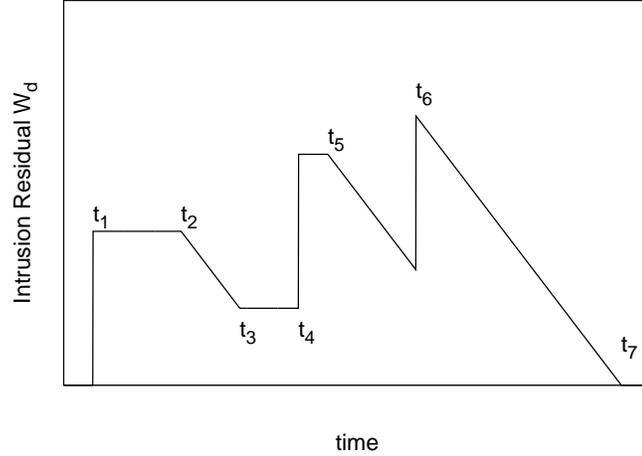
Figure 3.3: Illustration of intrusion residual function.

The function $W_d(t)$ helps us understand the intrusion behavior of the prob-ing traffic into $W(t)$. Before the arrival of probing packets, $W_d(t) = 0$. It gets an immediate increment of $s/C$ upon every probing packet arrival, where $s$ is the packet size. In $W(t)$'s busy periods without additional probing packet arrival, $W_d(t)$ remains unchanged. In $W(t)$'s idle periods without additional probing packet arrival, $W_d(t)$ deceases linearly with slope $-1$. Function $W_d(t)$ is monotonically non-increasing between every two adjacent probing packet ar-rivals. Figure 3.3 illustrates this behavior, where $(t_1, t_2)$ and $(t_3, t_5)$ are two busy periods in $W(t)$, and $(t_2, t_3)$ and $(t_5, t_7)$ are two idle periods in $W(t)$. Times $t_1$, $t_4$ and $t_6$ are the instants of probing packet arrivals. Time $t_7$ is the end point of the intrusive range.

Based on the above observations of $W_d(t)$, we state the following lemma without proof:

**Lemma 4** *When $W(t)$ is probed by a single packet $p$ of size $s$ arriving into the*

*hop at time $t_0$,*

$$W_d(t) = \begin{cases} 0 & t < t_0 \\ \max\left(0, \dfrac{s}{C} - I(t_0, t)\right) & t \geq t_0 \end{cases}. \qquad (3.28)$$

When $W(t)$ is probed by a packet train $\langle a_1, g_I, s, n \rangle$, we are often interested in computing

$$R_i(a_1) = W_d^-(a_i) = W_d^-(a_1 + (i-1)g_I) \qquad (3.29)$$

for $i = 1, 2, \ldots, n$. Metric $R_i(a_1)^3$ is the intrusion residual *caused* by the first $i - 1$ packets in the probing train $\langle a_1, g_I, s, n \rangle$ and *experienced* by packet $p_i$. In other words, the queuing delay of $p_i$ in the hop is given by:

$$\begin{aligned} \tilde{W}^-(a_i) &= W(a_i) + W_d^-(a_i) \\ &= W(a_i) + R_i(a_1). \end{aligned} \qquad (3.30)$$

The total sojourn time of $p_i$ at the hop is the sum of its service time and its queuing delay:

$$d_i - a_i = W(a_i) + R_i(a_1) + \frac{s}{C} \qquad (3.31)$$

As a direct result of Lemma 4, $R_i$ can be recursively computed as follows:

$$R_i = \begin{cases} 0 & i = 1 \\ \max\left(0, \dfrac{s}{C} + R_{i-1} - I(a_{i-1}, a_i)\right) & i > 1 \end{cases}. \qquad (3.32)$$

Denoting $s/C - I(a_{i-1}, a_i)$ by $y_i$, the second part of equation (3.32) can be expanded to the following non-recursive form:

$$R_i = \max\left(0, y_{i-1}, \sum_{k=i-2}^{i-1} y_k, \ldots, \sum_{k=1}^{i-1} y_k\right). \qquad (3.33)$$

We next discuss the second part of traffic interaction.

---

[3] When $a_1$ is irrelevant, we often write $R_i(a_1)$ as $R_i$.

### 3.2.3   Output Gaps of Individual Probing Trains

We first present a corollary. It is due to the work-conserving assumption. It says that the whole duration of any packet's stay at the hop is a hop busy period.

**Corollary 1** *For any packet arriving into the hop at time $t$ and departing from the hop at time $t + \delta$, $[t, t + \delta]$ is a hop busy period.*

Our next lemma describes the relationship between probing input and output for an individual packet train. It is the corner stone of our probing analysis. Previous work only revealed this result under certain conditions [15], [32]. The full picture, although simple and important, has remained undocumented.

**Lemma 5** *Assuming $\delta = (n-1)g_I$ and $W(t)$ is probed by a packet train $\langle a_1, g_I, s, n \rangle$, the output gap $g_O$ can be expressed as:*

$$
\begin{aligned}
g_O &= \frac{Y_\delta(a_1)g_I}{C} + \frac{s}{C} + \frac{\tilde{I}(a_1, a_n)}{n-1} \\
&= g_I + \frac{D_\delta(a_1)}{n-1} + \frac{R_n(a_1)}{n-1}.
\end{aligned}
\tag{3.34}
$$

**Proof:**   Examine hop activity of $\tilde{W}(t)$ within the time interval $[d_1, d_n]$. Notice that $(n-1)s/C$ time units are spent on serving all probing packets except $p_1$ and that

$$
\frac{V(a_n) - V(a_1)}{C} = \frac{Y_\delta(a_1)(n-1)g_I}{C} = \frac{Y_\delta(a_1)\delta}{C}
\tag{3.35}
$$

time units are spent on serving the cross traffic that has arrived at the hop during the time interval $[a_1, a_n]$. Thus the total hop working time in $[d_1, d_n]$ is given by

$$
\frac{Y_\delta(a_1)\delta}{C} + \frac{(n-1)s}{C}.
\tag{3.36}
$$

Also notice that $\tilde{I}(d_1, d_n)$ is the total idle time of the hop during this time interval. Since the sum of the hop working time in (3.36) and hop idle time must be equal to $d_n - d_1$, we immediately have the following:

$$d_n - d_1 = \frac{(n-1)g_I Y_\delta(a_1)}{C} + \frac{(n-1)s}{C} + \tilde{I}(d_1, d_n), \tag{3.37}$$

which leads to:

$$g_O = \frac{d_n - d_1}{n-1} = \frac{g_I Y_\delta(a_1)}{C} + \frac{s}{C} + \frac{\tilde{I}(d_1, d_n)}{n-1}. \tag{3.38}$$

Further, due to corollary 1, we get:

$$\tilde{I}(d_1, d_n) = \tilde{I}(a_1, a_n). \tag{3.39}$$

Substitute (3.39) back to (3.38), we proved the first equality in (3.34). For the second equality in (3.34), first recall from (3.31) that:

$$d_k = a_k + R_k(a_1) + W(a_k) + \frac{s}{C}, \quad k = 1, 2, \ldots, n. \tag{3.40}$$

Thus,

$$d_n - d_1 = (a_n - a_1) + R_n(a_1) + D_\delta(a_1). \tag{3.41}$$

Dividing both sides of (3.41) by $n - 1$, we get:

$$g_O = \frac{d_n - d_1}{n-1} = g_I + \frac{D_\delta(a_1)}{n-1} + \frac{R_n(a_1)}{n-1}. \tag{3.42}$$

This proved the second equality in (3.34). ∎

Lemma 5 shows that the output gap carries the information about $Y_\delta(a_1)$, which is potentially useful in cross-traffic measurements. However, the output gap is also contaminated by the noise information of $D_\delta(a_1)$, $\tilde{I}(a_1, a_n)$, and $R_n(a_1)$. In Lemma 2, we established the zero-mean nature for the first noise

term. The other two terms can have *positive* means in bursty cross-traffic. That is exactly where the response deviation comes from, as we show later. Meanwhile, we examine several useful bounds for these two terms.

From (3.32), noticing that $I(a_{i-1}, a_i)$ is no less than zero and applying mathematical induction to $i$, we get $0 \leq R_n \leq (n-1)s/C$. Combining with Lemma 5, we have:

**Corollary 2** *Again assuming $\delta = g_I(n-1)$, the following inequalities hold:*

$$\frac{D_\delta(a_1)}{n-1} + g_I \leq g_O \leq \frac{D_\delta(a_1)}{n-1} + g_I + \frac{s}{C}. \tag{3.43}$$

*The second inequality is tight iff $I(a_1, a_n) = 0$.*

Now we get into the second noise item $\tilde{I}(a_1, a_n)$. The next lemma leads to a bound for $\tilde{I}(a_1, a_n)$.

**Lemma 6** *For $k = 1, 2, \ldots, n-1$, we have:*

$$\begin{cases} \tilde{I}(a_k, a_{k+1}) = 0 & g_I \leq \dfrac{s}{C} \\ 0 \leq \tilde{I}(a_k, a_{k+1}) \leq g_I - \dfrac{s}{C} & g_I > \dfrac{s}{C} \end{cases}. \tag{3.44}$$

**Proof:** First, due to the probing intrusion behavior illustrated in Figure 3.3, we have:

$$\tilde{I}(a_k, a_{k+1}) = \max(0, I(a_k, a_{k+1}) - \frac{s}{C} - R_k), \tag{3.45}$$

where $R_k \geq 0$ and $0 \leq I(a_k, a_{k+1}) \leq (a_{k+1} - a_k) = g_I$. When $g_I \leq s/C$, $I(a_k, a_{k+1}) \leq s/C$. (3.45) becomes 0. Thus, the first part of (3.44) is proved.

When $g_I > s/C$, note that

$$\max(0, I(a_k, a_{k+1}) - \frac{s}{C} - R_k)$$
$$\leq \max(0, I(a_k, a_{k+1}) - \frac{s}{C})$$
$$\leq \max(0, g_I - \frac{s}{C}) = g_I - \frac{s}{C}. \tag{3.46}$$

This proves the second part of (3.44). ∎

Since the term $\tilde{I}(a_1, a_n)$ can be expressed as a sum:

$$\tilde{I}(a_1, a_n) = \sum_{k=1}^{n-1} \tilde{I}(a_k, a_k + 1), \tag{3.47}$$

we get the following bounds on the noise term $\tilde{I}(a_1, a_n)/(n-1)$ after combining (3.44) with (3.47):

$$\begin{cases} \dfrac{\tilde{I}(a_1, a_n)}{n-1} = 0 & g_I \leq \dfrac{s}{C} \\[2ex] 0 \leq \dfrac{\tilde{I}(a_1, a_n)}{n-1} \leq g_I - \dfrac{s}{C} & g_I > \dfrac{s}{C} \end{cases}. \tag{3.48}$$

Collecting Lemma 5 and (3.48), we get the following result.

**Corollary 3** *When $W(t)$ is probed by $\langle a_1, g_I, s, n \rangle$,*

$$\begin{cases} g_O = \dfrac{Y_\delta(a_1)g_I}{C} + \dfrac{s}{C} & g_I \leq \dfrac{s}{C} \\[2ex] \dfrac{Y_\delta(a_1)g_I}{C} + \dfrac{s}{C} \leq g_O \leq \dfrac{Y_\delta(a_1)g_I}{C} + g_I & g_I > \dfrac{s}{C} \end{cases}. \tag{3.49}$$

We call $(g_O C - s)/g_I$ the *intensity sampling estimator* (ISE). Corollary 3 implies that when ISE is used to estimate $Y_\delta(a_1)$, it is ensured to be correct only when $g_I \leq s/C$. When $g_I > s/C$, ISE's correctness is not guaranteed and it tends to overestimate $Y_\delta(a_1)$. The amount of overestimation, however, will not be more than $C - s/g_I$, as can be easily derived from the inequality in Corollary 3.

Finally, we must also notice an important relationship between $R_n(a_1)$ and $\tilde{I}(a_1, a_n)$. By subtracting the two expressions of $g_O$ in (3.34) and combining Lemma 3, we get:

$$\tilde{I}(a_1, a_n) = R_n(a_1) + I(a_1, a_n) - \frac{(n-1)s}{C}. \tag{3.50}$$

With the understanding of individual packet train probing, we are now in a position to derive the probing response curve.

## 3.3 Probing Response Curves

The probing response curve depends on a number of factors such as packet-train parameters, the inter-packet pattern, and cross-traffic characteristics. We assume a Poisson inter-probing pattern, because the asymptotic average of Poisson samples converges to the limiting time average of the sample-path being sampled. This property is known as PASTA (Poisson Arrivals See Time Averages) [41]. The average rate of Poisson sampling is assumed to be small enough so that the interference between adjacent trains can be neglected. We use $\langle \{T_m\}, g_I, s, n \rangle$ to denote a probing train series driven by a Poisson arrival process $\Lambda(t) = \max\{m \geq 0 : T_m \leq t\}$. We use $g_O^{(k)}$ to denote the output gap of the $k^{th}$ probing train $\langle T_k, g_I, s, n \rangle$ in the series, i.e., $g_O^{(k)} = (d_n^{(k)} - d_1^{(k)})/(n-1)$. The term $E[g_O]$ in (3.4) is defined as the limiting average of the discrete-time sample-path $g_O^{(k)}$:

$$E[g_O] = \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} g_O^{(k)}. \tag{3.51}$$

As mentioned, we use the notation of probability expectation to represent limiting time average, both for continuous-time sample-paths and for discrete-time sample-paths.We now introduce relevant concepts to characterize sample-path statistics and formally state a simplified sample-path version of PASTA that we use in subsequent derivations.

### 3.3.1 Frequency distribution and PASTA

**Definition 8** *For continuous-time sample-path $X(t)$, define indicator function $\Psi(x,t)$:*

$$\Psi(x,t) = \begin{cases} 1 & X(t) \leq x \\ 0 & X(t) > x \end{cases}. \tag{3.52}$$

*The frequency distribution function $P(x)$ of $X(t)$ is defined as following (assuming the limit exists for $\forall x$):*

$$P(x) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \Psi(x, t) dt. \qquad (3.53)$$

*For discrete-time sample-path $X_n$, define indicator function as:*

$$\Psi(x, n) = \begin{cases} 1 & X_n \leq x \\ 0 & X_n > x \end{cases}. \qquad (3.54)$$

*The frequency distribution function $P(x)$ of $X_n$ is defined as following (assuming the limit exists for $\forall x$):*

$$P(x) = \lim_{k \to \infty} \frac{1}{k} \sum_{n=1}^k \Psi(x, n). \qquad (3.55)$$

For a sample-path of stochastic vector process $\vec{X}(t)$, we can similarly define its frequency distribution function $P(\vec{x})$. The only trick is to interpret the $\leq$ and $>$ inequality symbols in (3.52) and (3.54) as a relation for every corresponding component in the vector.

**Lemma 7** *Assuming that $\vec{X}(t)$ is a continuous-time sample-path with frequency distribution $P(\vec{x})$, $T_k$ is a Poisson arrival sample-path, then the discrete-time sample-path $\vec{X}(T_k)$ also has frequency distribution $P(\vec{x})$.*

Lemma 7 basically says that Poisson sampling sees the sample-path frequency distribution. Consequently, Poisson sampling also sees the sample-path time average, which is just the expectation of the sample-path frequency distribution. PASTA is a classic queuing theory result obtained in early 1980's. Rigorously speaking, PASTA requires an assumption called LAA (Lack of Anticipation Assumption) on the Poisson arrival process, and the result holds in

"almost surely" sense, instead of pathwise sense. Practically, the Poisson process governing packet train probing is mostly made *independent* of the cross-traffic arrival process, which is a condition much stronger than the LAA assumption. Hence, in Lemma 7, we avoid the technical rigor that has little practical implication.

### 3.3.2   Bounds

We now obtain upper and lower bounds on the gap response curve.

**Theorem 3** *When $W(t)$ is probed by a Poisson packet-train series $\langle \{T_m\}, g_I \leq s/C, s, n \rangle$, the following equality holds:*

$$E[g_O] = \frac{g_I \lambda}{C} + \frac{s}{C}. \tag{3.56}$$

**Proof:**   Let $\delta = (n-1)g_I$. Using Corollary 3, $g_I \leq \dfrac{s}{C}$ implies:

$$E[g_O] = E\left[\frac{g_I Y_\delta(T_m) + s}{C}\right] = \frac{g_I E[Y_\delta(T_m)] + s}{C}. \tag{3.57}$$

Since $\{T_m\}$ is driven by Poisson arrivals, due to the PASTA property, we have:

$$E[Y_\delta(T_m)] = E[Y_\delta(t)]. \tag{3.58}$$

Combining (3.57), (3.58), and Lemma 1, we get (3.56). ∎

Rearranging the result of Theorem 3, we get:

$$\lambda = \frac{E[g_O]C - s}{g_I} = E\left[\frac{g_O C - s}{g_I}\right], \tag{3.59}$$

which explains when and why ISE can form an unbiased estimator for traffic intensity and thus for the available bandwidth.

**Theorem 4** *When $W(t)$ is probed by Poisson packet-train series $\langle\{T_m\}, g_I > s/C, s, n\rangle$, the following holds:*

$$\max\left(\frac{g_I\lambda + s}{C}, g_I\right) \leq E[g_O] \leq \min\left(g_I(1 + \frac{\lambda}{C}), g_I + \frac{s}{C}\right).$$

**Proof:**  Notice that when $g_I > s/C$:

$$E[g_O] \geq \frac{g_I E[Y_\delta(T_k)] + s}{C} = \frac{g_I E[Y_\delta(t)] + s}{C} = \frac{g_I\lambda + s}{C}. \tag{3.60}$$

Similarly, due to Corollary 2, PASTA, and Lemma 2, we have:

$$E[g_O] \geq g_I + \frac{E[D_\delta(T_k)]}{n-1} = g_I + \frac{E[D_\delta(t)]}{n-1} = g_I. \tag{3.61}$$

Collecting (3.60) and (3.61), we get:

$$\max\left(\frac{g_I\lambda + s}{C}, g_I\right) \leq E[g_O]. \tag{3.62}$$

For the upper bound of $E[g_O]$, first, from Corollary 3, PASTA, and Lemma 1, we get:

$$
\begin{aligned}
E[g_O] &\leq g_I\left(1 + \frac{E[Y_\delta(T_k)]}{C}\right) \\
&= g_I\left(1 + \frac{E[Y_\delta(t)]}{C}\right) = g_I\left(1 + \frac{\lambda}{C}\right).
\end{aligned} \tag{3.63}
$$

Then from Corollary 2, PASTA, and Lemma 2, we get:

$$
\begin{aligned}
E[g_O] &\leq \frac{E[D_\delta(T_k)]}{n-1} + \frac{s}{C} + g_I \\
&= \frac{E[D_\delta(t)]}{n-1} + \frac{s}{C} + g_I = g_I + \frac{s}{C}.
\end{aligned} \tag{3.64}
$$

Combining (3.63) and (3.64), we get:

$$E[g_O] \leq \min\left(g_I(1 + \frac{\lambda}{C}), g_I + \frac{s}{C}\right). \tag{3.65}$$

This concludes the proof of this theorem.  ∎

Theorem 4 provides both a lower bound and an upper bound for $E[g_O]$ when $g_I > s/C$. Combining the case when $g_I \le s/C$ as is stated in Theorem 3, we get the lower bound on $E[g_O]$ for the entire probing range $0 < g_I < \infty$ as follows[4]:

$$
L(E[g_O]) = 
\begin{cases}
\max\left(\dfrac{g_I \lambda + s}{C}, g_I\right) & g_I > \dfrac{s}{C} \\
\dfrac{s + g_I \lambda}{C} & g_I \le \dfrac{s}{C}
\end{cases}
$$

$$
= 
\begin{cases}
g_I & g_I > \dfrac{s}{C - \lambda} \\
\dfrac{s + g_I \lambda}{C} & g_I \le \dfrac{s}{C - \lambda}
\end{cases}
\quad (3.66)
$$

That is exactly model (3.4) we are trying to validate. However, Theorem 4 shows that (3.4) is a *lower bound* of $E[g_O]$, which does not necessarily equal to $E[g_O]$. Likewise, combining Theorems 3 and 4, we have the entire upper bound summarized as follows:

$$
U(E[g_O]) = 
\begin{cases}
\min\left(g_I(1 + \dfrac{\lambda}{C}), g_I + \dfrac{s}{C}\right) & g_I > \dfrac{s}{C} \\
\dfrac{s + g_I \lambda}{C} & g_I \le \dfrac{s}{C}
\end{cases}
$$

$$
= 
\begin{cases}
\dfrac{s}{C} + \dfrac{g_I \lambda}{C} & g_I \le \dfrac{s}{C} \\
g_I + \dfrac{g_I \lambda}{C} & \dfrac{s}{C} \le g_I \le \dfrac{s}{\lambda} \\
g_I + \dfrac{s}{C} & g_I \ge \dfrac{s}{\lambda}
\end{cases}
\quad (3.67)
$$

The real gap response curve is contained between these two bounds. We define the "response deviation" $\beta(g_I, s, n)$ as the difference between the real gap response curve and the lower bound given by (3.66). It can be expressed as following due to Theorem 4, Lemma 5, and PASTA:

$$
\beta(g_I, s, n) = 
\begin{cases}
\dfrac{E[\tilde{I}(t, t + (n-1)g_I)]}{n - 1} & g_I \le \dfrac{s}{C - \lambda} \\
\dfrac{1}{n - 1} E[R_n(t)] & g_I \ge \dfrac{s}{C - \lambda}
\end{cases}
\quad (3.68)
$$

---

[4]$L(.)$ and $U(.)$ denote lower bound and upper bound of a function respectively.

We next give a closed-form expression for the response deviation and thus for the probing response curves.

### 3.3.3 Closed-from Expression

Assuming $\delta = g_I$, note that both $R_n(t)$ and $\tilde{I}(t, t + (n-1)\delta)$ can be expressed as *deterministic* functions of an $(n-1)$-dimensional vector

$$\vec{B}_\delta^{(n-1)}(t) = \begin{pmatrix} B_\delta(t) \\ B_\delta(t + \delta) \\ ... \\ B_\delta(t + (n-2)\delta) \end{pmatrix}. \tag{3.69}$$

The exact functional expressions, on the other hand, are not very important at this point. Therefore, we can introduce the following notation:

$$\tilde{I}(t, t + (n-1)\delta) = \varphi(\vec{B}_\delta^{(n-1)}(t)), \tag{3.70}$$

$$R_n(t) = \psi(\vec{B}_\delta^{(n-1)}(t)), \tag{3.71}$$

where $\varphi(.)$ and $\psi(.)$ are some vector functions. It then becomes apparent that the response deviation depends on the sample-path limiting frequency distribution of $\vec{B}_\delta^{(n-1)}(t)$. Denoting by $P_\delta^{(n-1)}(\vec{x})$ this distribution function, the response deviation can be expressed by the following vector integrals:

$$\beta(g_I, s, n) = \begin{cases} \dfrac{1}{n-1} \displaystyle\int_\Omega \varphi(\vec{x}) dP_\delta^{(n-1)}(\vec{x}) & g_I < \dfrac{s}{C - \lambda} \\ \dfrac{1}{n-1} \displaystyle\int_\Omega \psi(\vec{x}) dP_\delta^{(n-1)}(\vec{x}) & g_I \geq \dfrac{s}{C - \lambda} \end{cases},$$

where $\Omega$ is an $(n-1)$-dimensional cube $[0, C]^{n-1}$.

To better understand these results, we now consider a degenerated case where $n = 2$, i.e., the packet-pair probing case. Note that both $\varphi(\vec{x})$ and $\psi(\vec{x})$

(a) gap response deviation     (b) gap response curve     (c) rate response curve
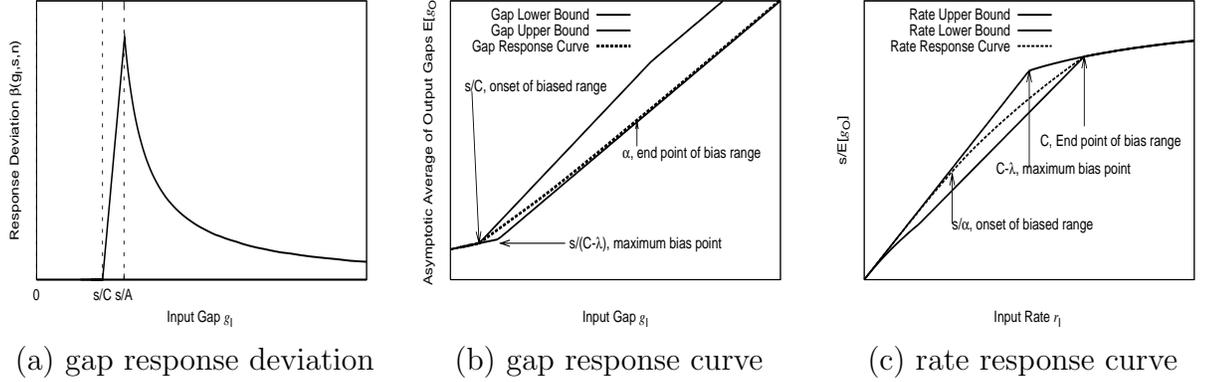
Figure 3.4: Illustrations of (a) the gap response deviation, (b) gap response curve, and (c) rate response curve in the entire input range.

become scalar functions and have simple expressions with respect to $B_\delta(t)$:

$$\tilde{I}_\delta(t) = \varphi(B_\delta(t)) = \max\left(0, \frac{B_\delta(t)\delta - s}{C}\right), \tag{3.72}$$

$$R_2(t) = \psi(B_\delta(t)) = \max\left(0, \frac{s - B_\delta(t)\delta}{C}\right). \tag{3.73}$$

Therefore, we have the following results for the packet-pair probing response curve.

**Theorem 5** *Assuming that $W(t)$ is probed by Poisson packet-pair series $\langle\{T_m\}, g_I, s, 2\rangle$, observation interval $\delta = g_I$, and the $\delta$-interval available bandwidth sample-path $B_\delta(t)$ has frequency distribution function $P_\delta(x)$, the following holds:*

$$
\begin{aligned}
E[g_O] &= \frac{g_I\lambda + s}{C} + \int_{s/\delta}^{C} \frac{x\delta - s}{C} dP_\delta(x) \\
&= g_I + \int_0^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x).
\end{aligned}
\tag{3.74}
$$

**Proof:** We only need to show the following:

$$E[\tilde{I}_\delta(t)] = \int_{s/\delta}^{C} \frac{x\delta - s}{C} dP_\delta(x), \tag{3.75}$$

$$E[R_2(t)] = \int_0^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x). \tag{3.76}$$

Then combining Lemma 5, Lemma 1, Lemma 2, Lemma 7, and both equations above, we immediately get the theorem.

To prove (3.75), simply recall (3.72) and we have:

$$
\begin{aligned}
E[\tilde{I}_\delta(t)] &= E\left[\max\left(0, \frac{B_\delta(t)\delta - s}{C}\right)\right] \\
&= \int_{s/\delta}^{C} \frac{x\delta - s}{C} dP_\delta(x).
\end{aligned}
$$

For the second part, recall (3.73) and we have:

$$
\begin{aligned}
E[R_2(t)] &= E\left[\max\left(0, \frac{s - B_\delta(t)\delta}{C}\right)\right] \\
&= \int_{0}^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x).
\end{aligned}
$$

This proved the theorem. ∎

It immediately follows that the packet-pair response deviation is as following (where $g_I = \delta$):

$$
\beta(g_I, s, 2) = \begin{cases} \int_{s/\delta}^{C} \frac{x\delta - s}{C} dP_\delta(x) & g_I < \frac{s}{C - \lambda} \\ \int_{0}^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x) & g_I \geq \frac{s}{C - \lambda} \end{cases}. \tag{3.77}
$$

The response deviation is one of the previously unknown factors causing measurement errors in available bandwidth estimation techniques based on (3.4). Our closed-from expressions show that the response deviation is exclusively decided by the packet-train parameters and the available bandwidth sample-path distribution. Next, we show the full picture of the response curves for both the gap version and the rate version.

### 3.3.4 Full Picture

We now investigate the relationship between the response deviation given in (3.68) and the input gap $g_I$ while keeping all other parameters fixed. We first

present the results for the case of packet-pair probing.

**Theorem 6** *When $W(t)$ is probed by Poisson packet pair series $\langle\{T_m\}, g_I, s, 2\rangle$, the response deviation $\beta(g_I, s, 2)$ equals zero when input gap $g_I \in (0, s/C]$; it is a monotonically increasing function of $g_I$ in the input gap range $(s/C, s/(C-\lambda)]$; and it is a monotonically decreasing function of $g_I$ in the input gap range $(s/(C-\lambda), \infty)$. Furthermore, in the whole input gap range $(0, \infty)$, the response deviation is a continuous function of $g_I$. Finally, the response deviation $\beta(g_I, s, 2)$ monotonically converges to 0 as $g_I$ approaches infinity.*

**Proof:** When $g_I \in (0, s/C]$, $\beta(g_I, s, 2)$ equals to 0 due to Theorem 3. Next, we prove the continuity and monotonicity properties of $\beta(g_I, s, 2)$. Let $\delta = g_I$, we first show that $E[\tilde{I}_\delta(t)]$ is a continuous and monotonically increasing function of $\delta$ in the range $\delta \in (0, \infty)$. First, note for any $0 < \Delta$ and $t$, we have:

$$0 \le \tilde{I}_{\delta+\Delta}(t) - \tilde{I}_\delta(t) \le \Delta. \tag{3.78}$$

This difference defines a new sample-path, and we can compute its time average as follows:

$$0 \le E[\tilde{I}_{\delta+\Delta}(t) - \tilde{I}_\delta(t)] \le \Delta, \tag{3.79}$$

which can be rewritten as:

$$0 \le E[\tilde{I}_{\delta+\Delta}(t)] - E[\tilde{I}_\delta(t)] \le \Delta. \tag{3.80}$$

This already proves the monotonicity of $E[\tilde{I}_\delta(t)]$ with respect to $\delta$. Further, by taking the limit of (3.80) when $\Delta \to 0$, we have:

$$\lim_{\Delta \to 0} (E[\tilde{I}_{\delta+\Delta}(t)] - E[\tilde{I}_\delta(t)]) = 0. \tag{3.81}$$

This proves the continuity of $E[\tilde{I}_\delta(t)]$ with respect to $\delta$ in the range $(0, \infty)$. Similarly, we can prove the continuity and monotonic decreasing property of

$E[R_2(t)]$ with respect to $\delta$ in the range $(0, \infty)$. Combining the monotonicity properties for both $E[R_2(t)]$ and $E[\tilde{I}_\delta(t)]$, further recall (3.68), we proved the monotonicity properties of $\beta(g_I, s, 2)$ described in this theorem.

For continuity of $\beta(g_I, s, 2)$, note that when $\delta = s/(C - \delta)$, the equality $E[R_2(t)] = E[\tilde{I}_\delta(t)]$ can be easily obtained from Theorem 5. Combining this result with the continuity of $E[\tilde{I}_\delta(t)]$ and $E[R_2(t)]$, we proved the continuity of $\beta(g_I, s, 2)$ with respect to $g_I$ in the entire probing range.

We next prove the asymptotic property of $\beta(g_I, s, 2)$ as $g_I \to \infty$. First, note that due to (3.77), we have:

$$
\begin{aligned}
\lim_{g_I \to \infty} \beta(g_I, s, 2) &= \lim_{\delta \to \infty} \int_0^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x) \\
&= \lim_{\delta \to \infty} \left( \int_0^{s/\delta} \frac{s}{C} dP_\delta(x) - \int_0^{s/\delta} \frac{x\delta}{C} dP_\delta(x) \right) \\
&= \lim_{\delta \to \infty} \int_0^{s/\delta} \frac{s}{C} dP_\delta(x) - \lim_{\delta \to \infty} \int_0^{s/\delta} \frac{x\delta}{C} dP_\delta(x).
\end{aligned}
\tag{3.82}
$$

Note that the first item in (3.82) is zero:

$$
\lim_{\delta \to \infty} \int_0^{s/\delta} \frac{s}{C} dP_\delta(x) = \lim_{\delta \to \infty} \frac{s}{C} P_\delta(\frac{s}{\delta}) = 0,
$$

and the second item in (3.82) is also zero:

$$
\begin{aligned}
0 \le \lim_{\delta \to \infty} \int_0^{s/\delta} \frac{x\delta}{C} dP_\delta(x) &< \lim_{\delta \to \infty} \int_0^{s/\delta} \frac{\frac{s}{\delta}\delta}{C} dP_\delta(x) \\
&= \lim_{\delta \to \infty} \frac{s}{C} P_\delta(\frac{s}{\delta}) = 0.
\end{aligned}
$$

Hence, the limit of $\beta(g_I, s, 2)$ when $g_I \to \infty$ is zero. This concludes the whole proof. ∎

Packet-pair response deviation has very nice functional properties in terms of continuity and monotonicity. The response deviation $\beta(g_I, s, 2)$ is a hill-shaped function with respect to $g_I$ as shown in Figure 3.4(a), where it reaches

its maximum when $g_I = s/(C - \lambda)$. Our next theorem presents an inequality that links the packet-train and packet-pair response deviation.

**Theorem 7** *For any $n \geq 2$, the following holds:*

$$\frac{1}{n-1}\beta((n-1)g_I, (n-1)s, 2) \leq \beta(g_I, s, n) \leq \beta(g_I, s, 2).$$

**Proof:** We prove the theorem when $g_I \leq s/(C - \lambda)$. The proof when $g_I \geq s/(C - \lambda)$ is very similar and we omit it. First, we slightly refine our notations. we use $\tilde{I}(t_0, t_1, t, g_I, s, n)$ to denote $\tilde{I}(t_0, t_1)$ when the hop is probed by a single packet train $\langle t, g_I, s, n \rangle$. When $t = t_0$, we omit the third parameter and only write $\tilde{I}(t_0, t_1, g_I, s, n)$. We now prove the first part $\frac{1}{n-1}\beta((n-1)g_I, (n-1)s, 2) \leq \beta(g_I, s, n)$. Note that:

$$\beta((n-1)g_I, (n-1)s, 2)$$
$$= E[\tilde{I}(t, t+(n-1)g_I, (n-1)g_I, (n-1)s, 2)] \tag{3.83}$$

$$\beta(g_I, s, n)$$
$$= \frac{1}{n-1}E[\tilde{I}(t, t+(n-1)g_I, g_I, s, n)]. \tag{3.84}$$

The idle time in (3.83) can be expanded as:

$$\tilde{I}(t, t+(n-1)g_I, (n-1)g_I, (n-1)s, 2)$$
$$= \max(0, I(t, t+(n-1)g_I) - \frac{(n-1)s}{C}). \tag{3.85}$$

Due to (3.50), the idle time in (3.84) can be expanded as:

$$\tilde{I}(t, t+(n-1)g_I, g_I, s, n)$$
$$= I(t, t+(n-1)g_I) - \frac{(n-1)s}{C} + R_n(t). \tag{3.86}$$

Combining (3.85) and (3.86), further noticing that $R_n(t) \geq 0$, we have for $\forall t$,

$$\tilde{I}(t, t+(n-1)g_I, (n-1)g_I, (n-1)s, 2)$$
$$\leq \tilde{I}(t, t+(n-1)g_I, g_I, s, n). \tag{3.87}$$

This leads to

$$E[\tilde{I}(t, t + (n-1)g_I, (n-1)g_I, (n-1)s, 2)]$$
$$\leq E[\tilde{I}(t, t + (n-1)g_I, g_I, s, n)]. \tag{3.88}$$

Dividing both sides of (3.88) by $n - 1$, we get:

$$\frac{1}{n-1}\beta((n-1)g_I, (n-1)s, 2) \leq \beta(g_I, s, n). \tag{3.89}$$

Next we prove the second part $\beta(g_I, s, n) \leq \beta(g_I, s, 2)$. Notice that for $k = 0, 1, \ldots, n-2$,

$$\tilde{I}(t + kg_I, t + (k+1)g_I, g_I, s, 2)$$
$$= \max\left(0, I(t + kg_I, t + (k+1)g_I) - \frac{s}{C}\right), \tag{3.90}$$
$$\tilde{I}(t + kg_I, t + (k+1)g_I, t, g_I, s, n)$$
$$= \max\left(0, I(t + kg_I, t + (k+1)g_I) - \frac{s}{C} - R_k(t)\right). \tag{3.91}$$

Combining (3.90) and(3.91), noticing that $R_k(t) \geq 0$, we get:

$$\tilde{I}(t + kg_I, t + (k+1)g_I, t, g_I, s, n)$$
$$\leq \tilde{I}(t + kg_I, t + (k+1)g_I, g_I, s, 2). \tag{3.92}$$

This inequality also holds when we sum up all idle time from $k = 0$ to $k = n-2$:

$$\sum_{k=0}^{n-2} \tilde{I}(t + kg_I, t + (k+1)g_I, t, g_I, s, n)$$
$$\leq \sum_{k=0}^{n-2} \tilde{I}(t + kg_I, t + (k+1)g_I, g_I, s, 2). \tag{3.93}$$

Note that the left-side item in (3.93) is actually $\tilde{I}(t, t + (n-1)g_I, g_I, s, n)$. Computing the limiting time averages of both sides in (3.93), we get:

$$
\begin{aligned}
&E[\tilde{I}(t, t + (n-1)g_I, g_I, s, n) \\
&\leq \sum_{k=0}^{n-2} E[\tilde{I}(t + kg_I, t + (k+1)g_I, g_I, s, 2)] \\
&= (n-1)E[\tilde{I}(t, t + g_I, g_I, s, 2)].
\end{aligned}
\tag{3.94}
$$

Dividing $(n-1)$ at both sides of (3.94), we get:

$$
\beta(g_I, s, n) \leq \beta(g_I, s, 2).
\tag{3.95}
$$

This proves the second inequality in this theorem. ∎

This result tells us that the packet-train response deviation $\beta(g_I, s, n)$ has similar hill-shaped evolving trend with respect to $g_I$ since it is both lower-bounded and upper-bounded by hill-shaped functions. We conjecture that it is also continuous and has similar monotonicity properties described in Theorem 6.

In summary, the response deviation is significant only in the middle part of the whole probing range. We call that range the *deviated probing range.* The full picture of the gap response curve is illustrated in Figure 3.4(b). The whole probing range $(0, \infty)$ is divided into three segments. Interval $(0, s/C]$ is an undeviated region where the ISE formula $(Cg_O - s)/g_I$ forms an unbiased intensity estimator for $\lambda$. Region $(s/C, \alpha)$ is a deviated region where $E[g_O]$ is larger than what is given in (3.4), but smaller than the upper bound in (3.67) and the ISE formula overestimates $\lambda$. Finally, interval $(\alpha, \infty)$ is the second undeviated probing range where $E[g_O] = g_I$. Theoretically, this range often does not exist due to infinite $\alpha$. Practically, a sufficiently small deviation is taken as none. The probing point $s/(C - \lambda)$, associated with available bandwidth, is

the point where the response deviation is maximized and is not the same as the turning point $\alpha$. Further note that the upper bound of gap response curve as given in (3.67) is actually not a tight bound.

It is often more informative to look at the rate version of the response curve rather than the gap version, because it has a direct association with our measurement interests: traffic intensity and available bandwidth. Transforming (3.4) into the corresponding rate version, we get the rate upper bound:

$$U\left(\frac{s}{E[g_O]}\right) = \begin{cases} r_I & 0 < r_I \leq C - \lambda \\ C\dfrac{r_I}{r_I + \lambda} & r_I > C - \lambda \end{cases}. \tag{3.96}$$

Although (3.96) looks similar to (2.1), they are in fact very different since $E[r_O] = E[s/g_O] \neq s/E[g_O]$ and $E[r_O]$ has a different behavior from that of $s/E[g_O]$. Our conclusions are meant for $s/E[g_O]$, not for $E[s/g_O]$. Although TOPP proposes (2.1) as its rationale, its actual implementation is however based on (3.96). It is important to clarify this confusion.

Transforming (3.67) gives us the rate lower bound as follows.

$$L\left(\frac{s}{E[g_O]}\right) = \begin{cases} \dfrac{r_I C}{r_I + C} & 0 < r_I \leq \lambda \\ \dfrac{r_I C}{\lambda + C} & \lambda < r_I \leq C \\ \dfrac{r_I C}{r_I + \lambda} & C < r_I \end{cases}. \tag{3.97}$$

As illustrated in Figure 3.4(c), along the vertical direction, the rate response curve appears between the two bounds given above. Along the horizontal direction, the curve shows one negatively deviated probing region sandwiched by two undeviated probing regions.

### 3.3.5   The Impact of Packet Train Parameters

We now examine the impact of probing packet size on response deviation. First, we consider the rate response curve of packet-pair probing. At any fixed input rate point $r < C - \lambda$, let $s \to \infty$. This causes the sampling interval $\delta = s/r$ approach to infinity proportionally. Recall (3.77), we have:

$$
\begin{aligned}
\beta\left(\frac{s}{r}, s, 2\right) &= \int_0^r \frac{s - x\delta}{C} dP_\delta(x) \\
&= \int_0^r \frac{r\delta - x\delta}{C} dP_\delta(x) = \frac{\delta}{C} \int_0^r (r - x) dP_\delta(x) \\
&= \frac{\delta}{C} \left( r \int_0^r dP_\delta(x) - \int_0^r x dP_\delta(x) \right)
\end{aligned}
\tag{3.98}
$$

Applying integration by parts, we get:

$$
\int_0^r x dP_\delta(x) = r P_\delta(r) - \int_0^r P_\delta(x) dx.
\tag{3.99}
$$

Substituting (3.99) back to (3.98), we get

$$
\beta\left(\frac{s}{r}, s, 2\right) = \frac{\delta}{C} \int_0^r P_\delta(x) dx.
\tag{3.100}
$$

From (3.100), we get a sufficient and necessary condition for packet-pair response deviation at input rate $r < A$ to vanish when $s \to \infty$:

$$
\lim_{\delta \to \infty} \delta \int_0^r P_\delta(x) dx = 0.
\tag{3.101}
$$

Similarly, for any input rate $r > A$, a sufficient and necessary condition for packet-pair response deviation to vanish is:

$$
\lim_{\delta \to \infty} \delta \left( C - r - \int_r^C P_\delta(x) dx \right) = 0.
\tag{3.102}
$$

These conditions require the cross-traffic not only exhibit decaying variance or gradually concentrating distribution when the observation interval $\delta$ becomes

large, but also show sufficient decaying speed. Our experiments show that cross-traffic often satisfies these properties. Hence, larger probing packet size usually implies less response deviation. The same conclusion also holds For packet train probing due to the following theorem.

**Theorem 8** *For any input probing rate $r$, If*

$$\lim_{s\to\infty} \beta\left(\frac{s}{r}, s, 2\right) = 0, \tag{3.103}$$

*then for packet train of any length, we have:*

$$\lim_{s\to\infty} \beta\left(\frac{s}{r}, s, n\right) = 0, \quad \forall n > 2. \tag{3.104}$$

**Proof:** Recall Theorem 7, we have:

$$\frac{1}{n-1}\beta\left(\frac{(n-1)s}{r}, (n-1)s, 2\right) \leq \beta\left(\frac{s}{r}, s, n\right) \leq \beta\left(\frac{s}{r}, s, 2\right).$$

Taking the limits of all three terms in the above inequality and noticing (3.103), we get:

$$0 \leq \lim_{s\to\to\infty} \beta\left(\frac{s}{r}, s, n\right) \leq 0. \tag{3.105}$$

Hence, $\lim_{s\to\infty} \beta(\frac{s}{r}, s, n) = 0$. This proves the theorem. ∎

As to the impact of packet train length $n$, (3.33) shows that $R_n$ depends on a partial sum of series of random variables $y_i, i = 1, 2, \ldots, n-1$ summed in the reverse order. This is a classic form in random walk theory [42], which deals with partial sums of *i.i.d* random variables. Although it is unlikely for $y_i = s/C - I(a_i, a_{i+1})$ to be *i.i.d*, we make this assumption to keep the derivations tractable and apply random walk theory to conceptually understand the impact of train length on response deviation. Using the response deviation expression in (3.68), random walk theory says that if $E[y_i] < 0$, which is the case when

$g_I > s/(C - \lambda)$, $R_n$ converges in distribution to a finite-mean random variable as $n \to \infty$:

$$\lim_{n\to\infty} E[R_n] < \infty. \tag{3.106}$$

Consequently,

$$\lim_{n\to\infty} \frac{E[R_n]}{n-1} = 0. \tag{3.107}$$

On the other hand, when $E[y_i] \geq 0$, as is the case when $g_I \leq s/(C - \lambda)$, $R_n$ grows unbounded with probability 1 as $n \to \infty$. Note the following relationship between $R_n$ and $\tilde{I}(a_1, a_n)$:

$$\tilde{I}(a_n, a_{n+1}) = \max\left(0, I(a_n, a_{n+1}) - \frac{s}{C} - R_n\right). \tag{3.108}$$

Thus, there is a random point $n_0$ such that $\tilde{I}(a_n, a_{n+1})$ becomes 0 if $n > n_0$. And this $n_0$ converges in distribution to a finite-mean random variable as $n \to \infty$, Thus we have

$$\lim_{n\to\infty} E[\tilde{I}(a_1, a_n)] < \infty, \tag{3.109}$$

$$\lim_{n\to\infty} \frac{1}{n-1} E[\tilde{I}(a_1, a_n)] = 0. \tag{3.110}$$

This explains why the response deviation can be overcome by long packet trains. Even when $y_i$ are not *i.i.d* random variables and the above argument does not fully apply, it at least tells us why the response deviation can be mitigated, which is quite non-intuitive.

### 3.3.6   Discussion

We now briefly mention how sensitive our results are with respect to the assumptions made in this chapter. First, notice that the simple traffic-arrival assumption is made solely to avoid getting into unnecessary technical details.

Even when batch arrivals are allowed, simple arrivals occur almost everywhere along the time axis, and all the results we obtained so far remain valid.

In this chapter, we also assumed infinite buffer space in the hop. Hence, our results are valid when buffer space is sufficiently large and packet loss can be neglected. In the case of otherwise, the equality $A = C - \lambda$ becomes invalid. The analysis of the impact of buffer size on bandwidth estimation requires future work.

We further assumed a Poisson inter-probing pattern. This can be relaxed to more general ASTA [25] sampling and as long as the sampling pattern has decent ASTA properties, all of our conclusions hold. In the case of non-negligible ASTA bias, most measurement techniques would fail and nothing interesting is left for discussion. ASTA bias is another source of measurement error that has never been studied or evaluated before. We consider it beyond the scope of the thesis.

Finally, we made two sample-path assumptions on cross-traffic and avoided assuming cross-traffic stationarity. The later however was an assumption commonly made in prior work. Our results are applicable to but not limited to stationary cross-traffic. More information regarding this issue is given in the appendix.

Next, we present our experimental methodology to compute the probing response curve and observe the response deviation quantitatively.

## 3.4   Experimental Results

To characterize the response deviation, we need to obtain the limiting averages of the probing output. In this section, we propose two experimental procedures

to compute the probing response curves with supervised precision. The first procedure is *period testing*, applicable to periodic traffic such as CBR. The second procedure is *trace-driven testing*, applicable to aperiodic traffic. We first apply the former to CBR traffic to verify our analytical results. We then apply the latter to several additional traffic traces to examine the relationship between response deviation and packet-train parameters.

### 3.4.1 Period Testing

The CBR (Constant Bit Rate) traffic we consider here is the one with a fixed packet size, fixed inter-packet delay, and periodical triangle-wave workload sample-path showed in Figure 3.2(b). In this thesis, a traffic is called bursty if its cumulative arrival sample-path $V(t)$ is not a linear function of $t$. Hence, all but constant-rate fluid traffic is bursty. In this light, CBR cross-traffic is arguably the simplest type of bursty traffic; however, it is also very important since we believe that any available bandwidth estimation technique must be shown accurate in CBR cross-traffic before being tested in more complex environments.

It is clear that CBR traffic satisfies both stability assumptions we made. Period testing on CBR traffic operates as follows. Assume a scenario with CBR cross-traffic packet size $s_c$, inter-packet delay $T$, hop capacity $C$, and $s_c/C < T$. Without loss of generality, we let the first packet arrive to the router at time instance 0. We divide the time interval $[0, T]$ into $m$ equal-size sub-intervals. For all $k = 0, 1, 2...m - 1$, we compute the output gap $g_O^{(k)}$ of the probing train $\langle T(2k+1)/2m, g_I, s, n \rangle$. The average metric $\sum_{k=0}^{m-1} g_O^{(k)}/m$ of the output gaps is used as an approximation of $E[g_O]$. The departure time of the last packet in the

probing train is calculated using (3.31), where $W(a_n)$ can be easily computed due to the periodicity of the CBR workload sample-path. Also note that $R_n(a_1)$ can be recursively computed using (3.32). Thus, period testing can be conducted using deterministic computation without the use of `ns2`.

The validity of period testing is due to the following theorem:

**Theorem 9** *Let $E[g_O]$ be the asymptotic average of output gaps when the hop is probed by Poisson packet train series $\langle \{T_m\}, g_I, s, n \rangle$. Let $g_O(t)$ be the output gap when the hop is probed by a single packet train $\langle t, g_I, s, n \rangle$. Assuming the workload sample-path $W(t)$ associated with cross-traffic is a periodic function in the time interval $(0, \infty)$ and $T$ is period duration, the following holds:*

$$E[g_O] = \frac{1}{T} \int_0^T g_O(u) du. \tag{3.111}$$

**Proof:** First notice that, due to the periodicity of $W(t)$, the associated sample-paths such as $D_\delta(t)$ and $I_\delta(t)$ also have the same periodicity for all $\delta > 0$. Recall Lemma 5, which says $g_O(t)$ is a deterministic function of those sample-paths. Thus, $g_O(t)$ is also periodic with period duration $T$. It immediately follows that:

$$\frac{1}{T} \int_0^T g_O(u) du = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g_O(u) du. \tag{3.112}$$

On the other hand, note that due to PASTA, we have:

$$E[g_O] = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g_O(u) du. \tag{3.113}$$

Combining (3.112) and (3.113), we proved the theorem. ∎

Period testing essentially approximates the right-side item in (3.111) using $\sum_{k=0}^{m-1} g_O(T(2k+1)/2m)/m$. This approximation can be made arbitrarily precise by choosing sufficiently large $m$. Next, we introduce two supervision strategies

(a) rate response curve    (b) relative rate response deviation
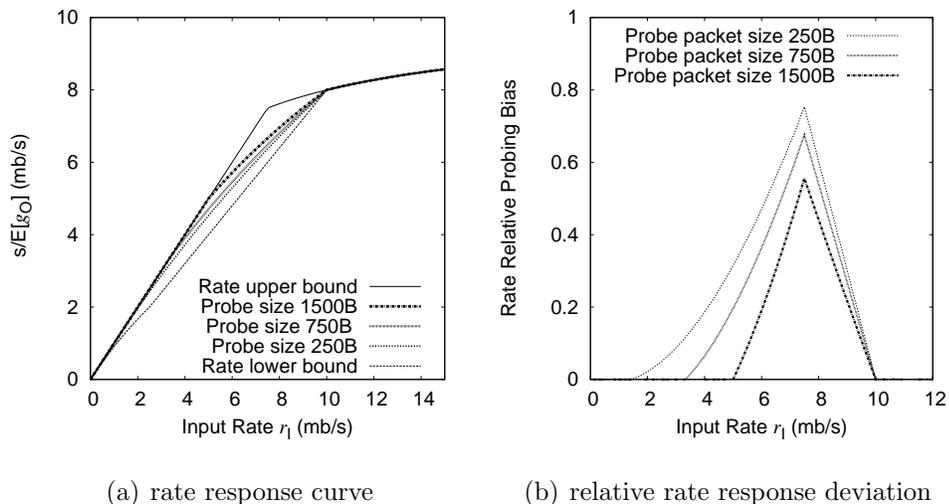
Figure 3.5: Packet pair probing in CBR cross-traffic: (a) Rate response curves, (b) relative rate response deviation. C=10mb/s, $\lambda$=2.5mb/s.

to help decide the number of samples $m$. Both are also applicable to trace-driven testing.

In the first method called *self supervision*, we iteratively double the number of samples and stop when there is little or no difference between the results produced in consecutive iterations. In the second method called *region supervision*, we make sure that $m$ is large enough so that the results of period testing are in agreement with those predicted by (3.4) or (3.96) in the undeviated probing range.

In our experiment, we choose $s_c = 1,500$ bytes, $C = 10$ mb/s, and $\lambda = 2.5$ mb/s. Thus, the inter-packet spacing of CBR cross-traffic is 4.8 ms. Using our supervision strategies, we find that 500 samples can provide very good precision and the results do not significantly differ from those obtained using $1,000$ or more samples.

Figure 3.5(a) shows the rate response curves when the hop is probed by

(a) gap response curve       (b) rate response curve
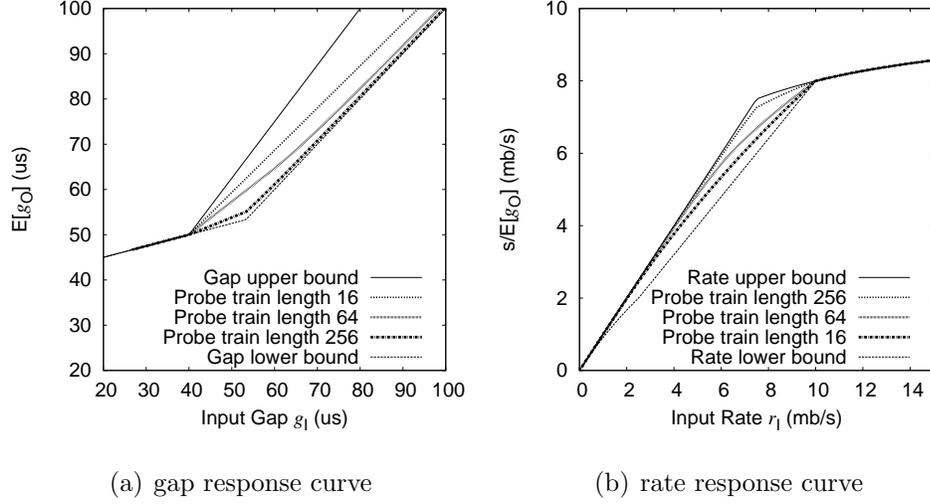
Figure 3.6: Packet train probing in CBR traffic: (a) Gap response curves, and (b) rate response curves. C=10mb/s, $\lambda$=2.5mb/s.

packet pairs. The legends are sorted in the same order as their corresponding curves appear vertically in the figure, and we do this whenever possible for all figures to make them easier to read.

Figure 3.5(b) shows the *relative* rate response deviation, defined as:

$$\frac{\min\left(r_I, \dfrac{r_I C}{\lambda + r_I}\right) - \dfrac{s}{E[g_O]}}{C - \lambda - \dfrac{(C - \lambda)C}{\lambda + C}}, \tag{3.114}$$

where the numerator is the absolute rate response deviation and the denominator is the difference between the rate upper bound and the rate lower bound when the input probing rate equals to the available bandwidth $C - \lambda$. This difference is an upper bound of absolute rate response deviation. Hence, the relative response deviation metric takes values in $[0, 1]$.

As shown in Figure 3.5, the response deviation is clearly noticeable for all three cases. The deviated regions are around (5 mb/s, 10 mb/s) for $s = 1500$ bytes, (3.5 mb/s, 10 mb/s) for $s = 750$ bytes, and (1.7 mb/s, 10 mb/s) for

$s = 250$ bytes. The relative response deviation (3.114) also exhibits high amplitude up to 0.5-0.8, meaning that, at certain probing ranges, the rate response curves are much closer to the lower bound than to the upper bound. Also note that as probing packet size $s$ increases, both the deviation range and deviation amplitude shrink. Further, the strongest deviation appears at the available bandwidth point for all three cases, which is 7.5 mb/s in our case. Finally, the response deviation appear monotonic at both sides of the available bandwidth probing point. These observations are in agreement with our theoretical findings.

Figure 3.6 shows gap and rate response curves when the hop is probed by packet trains. The probing packet size is 50 bytes. The reason why we use small probing packet size is to show that long trains can compensate for the deviation introduced by the small probing packet size. The figure shows the response curves for train lengths 16, 64, and 256 packets. From Figure 3.6, we see that the response deviation is clear, but diminishes as train length increases.

## 3.4.2  Trace-Driven Testing

**Traffic Traces**

In this section, we compare response deviation using four different cross-traffic types: CBR traffic, Poisson traffic with constant packet size (PCS), Poisson traffic with packet sizes (in bytes) uniformly distributed in $[1, 1500]$ (PUS), and Pareto *on/off* traffic (POF). Hop capacity $C$ is fixed at 10 mb/s. The cross-traffic packet size is 750 bytes for CBR, PCS, and the *on* period of POF. The average sending rate is 500 packets per second for CBR, PCS, and PUS. The mean duration of POF *on/off* periods is 10 and 5 ms, respectively. The Pareto

(a) Convergence delays       (b) Convergence errors
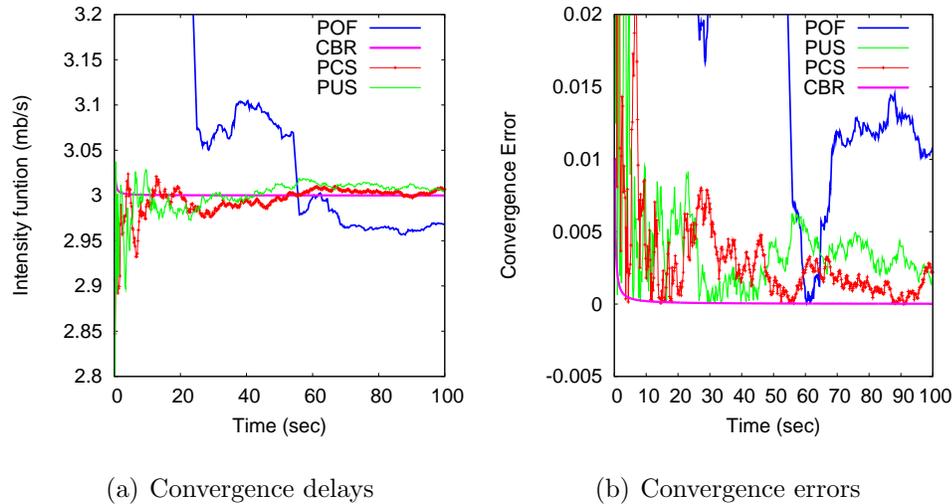
Figure 3.7: (a) Function $\mathcal{I}(t)$ shows the convergence delays, and (b) Function $\mathcal{R}(t)$ shows convergence errors for the four traffic traces.

shape parameter $\alpha$ for the duration of both *on/off* periods is set to 1.9 so that their variance is infinite. In POF *on* periods, the source sends CBR traffic at 750 packets per second. Given these settings, all four cross-traffic types have an average traffic intensity equal to 3 mb/s.

Since all but CBR traffic have aperiodic hop workload sample-path, we cannot apply period testing to obtain their response curves. Instead, we employ trace-driven testing to compute the response curves for the other three traffic types. We use RNGs (random number generators) to produce four packet-arrival traces, one for each traffic type. These traces record the time instances of all packet arrivals and their sizes within a period of 100 seconds. Before we explain how trace-driven testing works, we first show that these traffic traces satisfy the two cross-traffic stability assumptions we made.

In Figure 3.7(a), we plot function $\mathcal{I}(t) = V(t)/t$ for the four traffic traces. As shown in the figure, all traffic types exhibit *intensity* stability despite the

big differences in their convergence delays. Figure 3.7(b) shows the intensity convergence error defined as:

$$\mathcal{R}(t) = \frac{|\mathcal{I}(t) - 3 \text{ mb/s}|}{3 \text{ mb/s}}. \tag{3.115}$$

As demonstrated in Figure 3.7, CBR shows the fastest convergence speed. In about 10 seconds, CBR converges to the 0.2%-neighborhood of the limiting value, i.e., $\mathcal{R}(10) \leq 0.002$. PCS and PUS also converge relatively fast, but much slower than CBR. In 10 seconds, both PCS and PUS converge to the 1%-neighborhood of the desired 3 mb/s. PCS converges a little faster than PUS but the difference is small. POF shows the slowest convergence speed[5]. It reaches the 1.5%-neighborhood in about 60 seconds.

The four traffic traces also exhibit *workload* stability when they are injected in a hop of capacity $C = 10$ mb/s. This is theoretically provable. Using queueing theory, we can directly compute the limiting time average of the workload process for these four traffic types. The existence of workload limiting time average implies workload stability. More details are given in the appendix.

**Testing Procedure**

Trace-driven testing is grounded on the following corollary:

**Corollary 4** *Let $E[g_O]$ be the asymptotic average of output gaps when the hop is probed by Poisson packet train series $\langle \{T_m\}, g_I, s, n \rangle$. Let $g_O(t)$ be the output gap when the hop is probed by a single packet train $\langle t, g_I, s, n \rangle$. Then the following holds due to PASTA:*

$$E[g_O] = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g_O(u)du. \tag{3.116}$$

---

[5]POF carries some flavor of self-similar traffic.

Trace-driven testing essentially approximates the right-side item in (3.116) by computing the time average of $g_O(t)$ in a finite time interval $[0, t_0]$. The approximation can be made arbitrarily accurate when sufficiently large $t_0$ is used. We choose $t_0$ based on the convergence error function $\mathcal{R}(t)$ of the traffic traces, since a small value of $\mathcal{R}(t_0)$ is a good indication that the traffic statistics in $[0, t_0]$ has sufficiently converged to its equilibrium statistics.

Once $t_0$ is chosen, trace-driven testing computes the sample average $\sum_{k=0}^{m-1} g_O(t_0(2k+1)/2m)/m$ and uses it as an approximation of $\int_0^{t_0} g_O(u)du/t_0$, where $m$ is decided by the two supervision strategies discussed before. The computation of the output gap $g_O^{(k)}$ of the probing packet train $\langle t_0(2k+1)/2m, g_I, s, n \rangle$ again relies on (3.31) and (3.32), where the workload $W(t)$ at any time instance can be computed based on cross-traffic trace and hop capacity.

In our experiment, we choose $t_0 = 20$ seconds for PCS and PUS, which leads to $\mathcal{R}(t) \leq 0.01$, and $t_0 = 60$ seconds for POF, which ensures $\mathcal{R}(t) \leq 0.015$. For CBR, we still use period testing. In what follows, we first compute the response curves for several fixed packet-train parameters. We then study the impact of packet-train parameters on response deivation.

**Results for Fixed Packet-train Parameters**

Figure 3.8(a) shows the rate response curves for the four traces when the hop is probed using packet pairs. We computed the output rate $s/E[g_O]$ at 140 input rate points, from 1.0 mb/s to 14.0 mb/s with a 0.1 mb/s increment. We applied region supervision to decide the number of samples. That is, at each input rate in [10.0 mb/s, 14.0 mb/s], the number of samples is made large enough so that the output rate $s/E[g_O]$ computed in trace-driven testing is within the 1%-neighborhood of the value predicted by fluid model (3.96). This required

(a) Packet-pair probing          (b) Packet-train probing
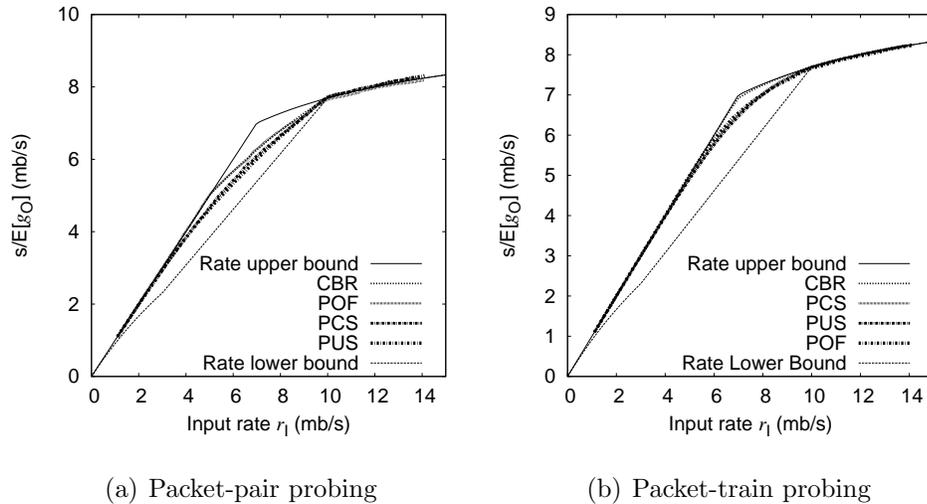
Figure 3.8: Rate response curve for the four cross-traffic traces: (a) probing pairs, (b) 16-packet trains (probing packet size 750 bytes).

500 samples for CBR, 1,000 samples for PCS and PUS, and 2,000 samples for POF.

As showed in Figure 3.8(a), the rate response curve of POF is virtually indistinguishable from that of CBR. The PCS and PUS curves are also very close to each other. However, it is interesting to note that the curve for POF is closer to rate upper bound than the curves for PUS and PCS, meaning that it suffers *less* response deviation. This indicates that, for fixed packet train parameters, cross-traffic of more burstiness does not necessarily imply larger response deivation. We explain the reasons in a short while.

Figure 3.8(b) shows the rate response curves for the four traces when the hop is probed using 16-packet trains. For the CBR trace, the response curve is almost undeviated and hardly distinguishable from the rate upper bound in the figure. The response deviations are still clear for the other three traces; and those three curves are very close to each other. This shows that, as the probing

(a) log-log plotting of (a)

(b) log-log plotting of (c)
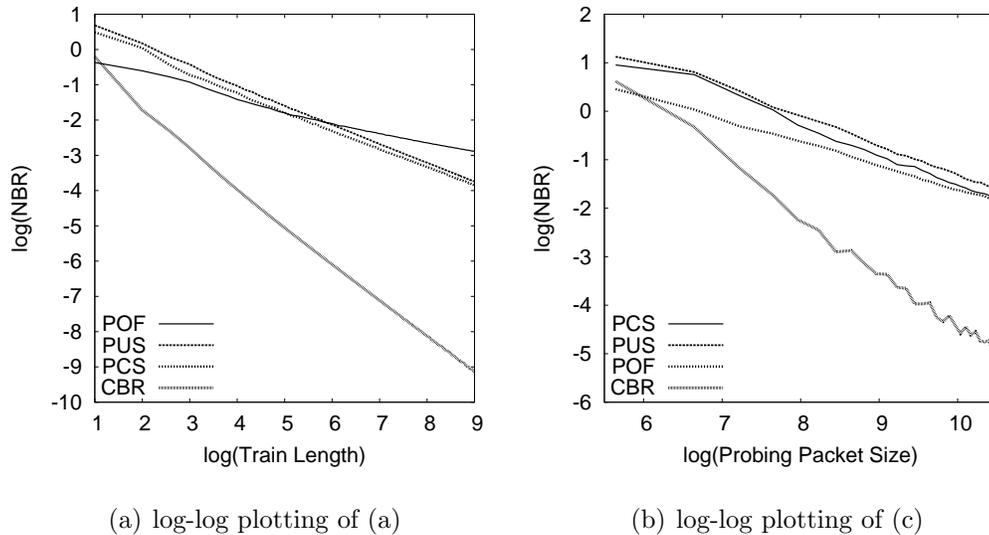
Figure 3.9: NBR for the four cross-traffic traces: (a) probing train length from 2 to 512. (b) log scale plotting of (a). (c) probing packet size from 50 bytes to 1500 bytes. (d) log scale plotting of (c).

train length increases, the response deivation diminishes. For cross-traffic of different burstiness, the diminishing rate is different. The response deivation for POF vanishes at a rate lower than those of the other three.

## Impact of Packet-train Parameters

Since we constantly observe that the response curves suffer the largest response deviation at the available bandwidth point, we define a metric called NBR (Normalized Bias Ratio) to characterize the amount of deviation in a rate response curve. Assuming $\mathbf{r}$ is the output rate $s/E[g_O]$ when the input rate is $A = C - \lambda$, we define:

$$\text{NBR} = \frac{A - \mathbf{r}}{\mathbf{r} - \dfrac{AC}{C + \lambda}}, \qquad (3.117)$$

which is the distance of the actual curve to its upper bound divided by the distance to its lower bound, given that the input probing rate is equal to the available bandwidth $A$. The NBR metric takes values in $[0, \infty)$, where larger NBR values indicate more response deviation in the response curve. We next investigate the relationship between NBR and packet-train parameters.

For all four traces, we computed NBR using probing packet sizes between 50 and 1500 bytes with 50-byte increasing step and probing train lengths between 2 and 512 packets with 2-packet increasing step. Thus, in total, we have $256 \times 30 = 7,680$ different packet-train parameters for each of the four traces. For each packet-train parameters, we calculate the output rate $\mathbf{r}$ in (3.117) using trace-driven testing with 2,000 samples.

Figure 3.9(a) shows NBR for the four traces using $s = 750$ bytes. In all four traces, NBR decreases as the probing train length increases and this relationship appears to be a power-law function as is confirmed by our log-log scale plotting in Figure 3.9(b). Figure 3.9(c) shows NBR when train length is fixed at 16 packets and the probing packet size varies from 50 bytes to 1500 bytes . We again observe a power-law decrease of NBR with respect to the increase in the probing packet size as showed in the log-scale plotting in Figure 3.9(d). Conjecturing that the relationship between NBR, probing size $s$, and train length $l$ can be modeled using function NBR $= k/(s^{\alpha_1} l^{\alpha_2})$, we get:

$$\log(\text{NBR}) = \log(k) - \alpha_1 \log(s) - \alpha_2 \log(l), \tag{3.118}$$

To obtain further insight into this formula, we plot 3D charts of NBR$(s, l)$ on a log-log scale for all four traces and indeed observed four flat planes. Figure 3.10 shows the four NBR planes.

We use 3D-fitting to find the parameters of the four planes. All least-square

log(NBR)

log(NBR)

(a) CBR plane

(b) PUS plane

log(NBR)

log(NBR)

(c) PCS plane

(d) POF plane
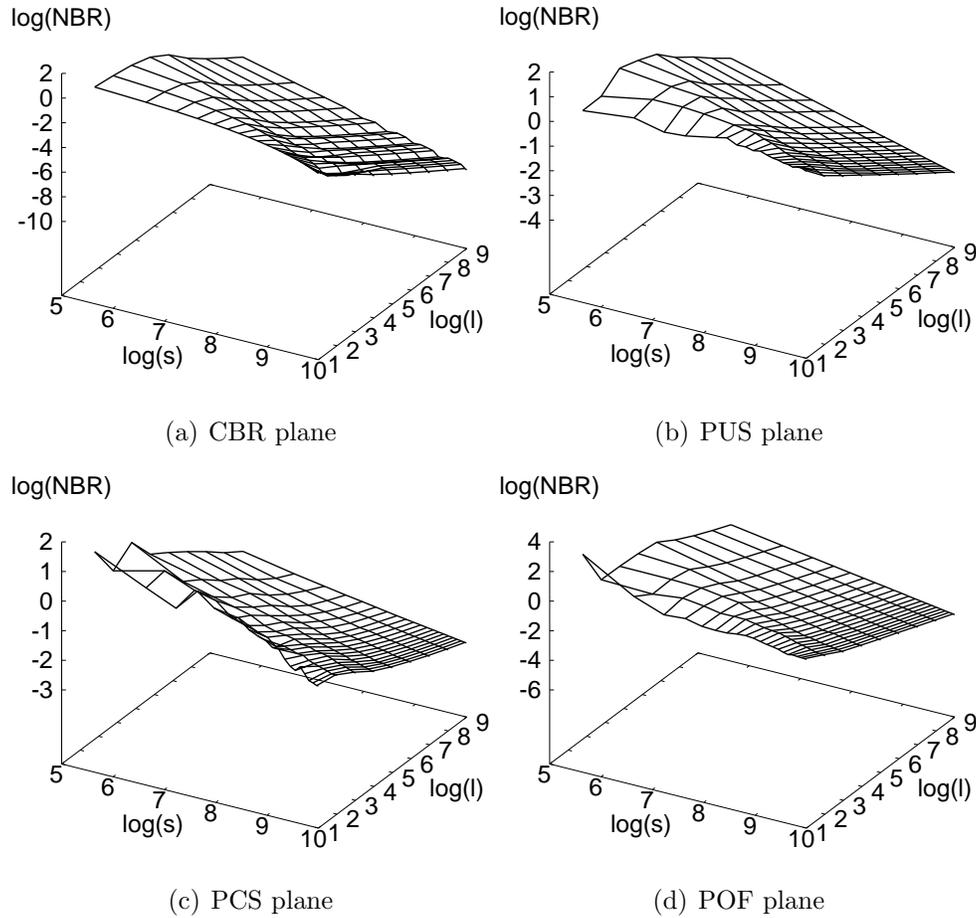
Figure 3.10: NBR$(s, l)$ for four types of cross-traffic on log-log scale.

fitting errors are less than 2%, indicating that the power-law function (3.118) is a reasonable model for NBR. Curve-fitting results are given in Table 3.1, which shows that traffic with more burstiness has smaller values of $\alpha_1$ and $\alpha_2$. This explains why the response deviation in POF is harder to overcome than those in the other three cross-traffic traces.

**Discussion**

The experimental results we obtained in trace-driven testing agree with our analytical findings very well. Furthermore, our results show that with fixed

Table 3.1: 3D-fitting results for NBR planes.

|  | $\alpha_1$ | $\alpha_2$ | $\log(k)$ |
|---|---|---|---|
| CBR | $1.103 \pm 0.017$ | $0.993 \pm 0.008$ | $10.53 \pm 0.175$ |
| PCS | $0.562 \pm 0.006$ | $0.534 \pm 0.003$ | $6.300 \pm 0.058$ |
| PUS | $0.524 \pm 0.008$ | $0.539 \pm 0.004$ | $6.111 \pm 0.077$ |
| POF | $0.413 \pm 0.007$ | $0.338 \pm 0.003$ | $4.000 \pm 0.074$ |

packet-train parameters, more cross-traffic burstiness does not necessarily implies more response deviation. This response deviation, however, is more difficult to overcome by increasing the probing packet size or probing train length.

To understand this phenomenon, recall that traffic burstiness relates to how fast the traffic becomes "smooth" with respect to the increase of observation intervals rather than how "smooth" the traffic appears given a fixed observation interval. Hence, it is usual that for a given observation interval, POF has smaller second order statistics than Poisson traffic and appears "smoother", leading to less response deviation when packet trains are constructed to sample the traffic in such an observation interval. As the train length or packet size increases, the observation interval increases, Poisson traffic becomes smooth quicker than POF. Therefore, the response deviation is also overcome quicker.

Even though we do not offer a precise interpretation for the power-law relation between NBR metric and packet-train parameters, we believe that it is related to the evolving trend of available bandwidth frequency distribution with respect to the increase of observation interval. This view is supported by the closed-form expression of response deviation, which shows that there is no other factor that can decide the NBR metric.

## 3.5    Implications

Among the five representative proposals TOPP, IGI/PTR, Spruce, `pahtload`, and `pathChirp`, the first three directly fall under the umbrella of our work. The last two techniques have quite a few tunable parameters and their behavior is complex. We will consider them in our future work.

### 3.5.1    TOPP

Figure 3.11 shows the rate response curves for the four traces when the hop is probed using $1,500$-byte packet pairs (as suggested in [28]). The curves are transformed using formula (2.3) so that TOPP can apply segmented linear regression to obtain the hop capacity and available bandwidth information. In the order of closeness to TOPP's expected piece-wise linear curve appear the response curves of CBR, POF, PCS and PUS. TOPP uses the second segment, assuming that it is the one with the hop information. However, the deviated probing range usually appears as the second segment unless it is very small and undetectable. In Figure 3.11, all the deviated ranges are very clear and will be incorrectly acted upon by TOPP. Table 3.2 shows the results of a linear regression applied to the deviated response curves according to the basic algorithm in TOPP. As the table shows, the available bandwidth is significantly *underestimated*, especially for PUS and PCS. Both the hop capacity and cross traffic intensity are significantly *overestimated*. To assure asymptotic accuracy, TOPP has to apply additional techniques to bypass these segments in the deviated probing range.
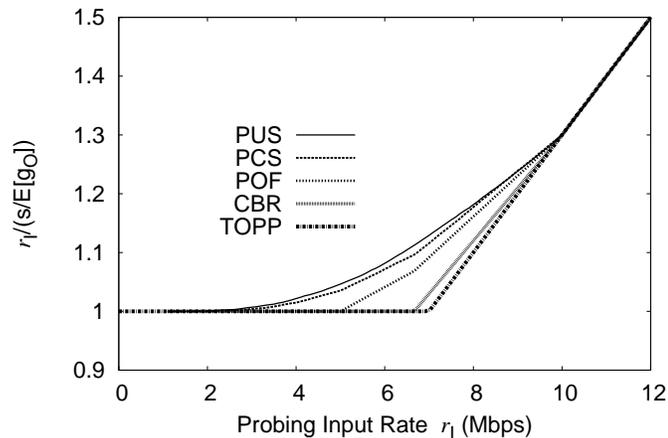
Figure 3.11: TOPP-transformed rate response curves.

Table 3.2: TOPP results (in mb/s) using the deviated segment (correct values: $C = 10$ mb/s, $A = 7$ mb/s).

|  | estimated $C$ | estimated $\lambda$ | estimated $A$ |
|---|---|---|---|
| CBR | 11.11 | 4.44 | 6.67 |
| PCS | 35.81 | 32.38 | 3.43 |
| PUS | 32.51 | 29.24 | 3.28 |
| POF | 23.38 | 18.36 | 5.02 |

## 3.5.2 IGI/PTR

PTR uses the probing output rate, $s/E[g_O]$, at the turning point to estimate the available bandwidth. As we established, the turning point usually is not the available bandwidth point. It can be associated with a rate much smaller than available bandwidth. Thus, theoretically-speaking, PTR is a *negatively biased* available bandwidth estimator in all single-hop paths.

As an estimator of cross-traffic intensity, the IGI formula

$$\lambda = E\left[\frac{\sum_{1 \leq i < n, d_{i+1} - d_i > g_I} C(d_{i+1} - d_i - \frac{s}{C})}{d_n - d_1}\right] \qquad (3.119)$$

is negatively biased when $g_I \leq s/C$. This is clear when comparing (3.119) with the ISE equation (3.59), which has the same numerator but smaller denominator than those of IGI. Recall that in [15], the IGI estimator is applied at the turning point where $a_n - a_1 = E[d_n - d_1]$. In that case, IGI has the same denominator, but a smaller numerator compared to ISE. According to Theorem 4, ISE is a positively-biased intensity estimator at the turning point, which suggests that IGI can be viewed as an estimator with a heuristical compensator for this bias. We use trace-driven testing to examine the performance of IGI's bias compensation. We use probing packet size 750 bytes and train length 64 packet as suggested in [15]. For comparison purposes, we also examine the ISE estimator and the PTR available bandwidth estimator.

Figure 3.12 shows these results for the four cross-traffic. The figure clearly shows that IGI provides a good estimate of cross-traffic intensity $\lambda$ at the available bandwidth point $A = 7$ mb/s, while not at the turning point $T \approx 6$ mb/s for all but CBR. When the input probing rate is small, IGI formula is not a converging estimator and the results are unstable.

For highly bursty traffic such as POF, the turning point falls into the unstable region and IGI does not ensure consistent results.

### 3.5.3 Spruce

Spruce uses ISE with input probing rate $C$ to estimate cross-traffic intensity. Thus, it is unbiased according to Theorem 3. Although this approach is more
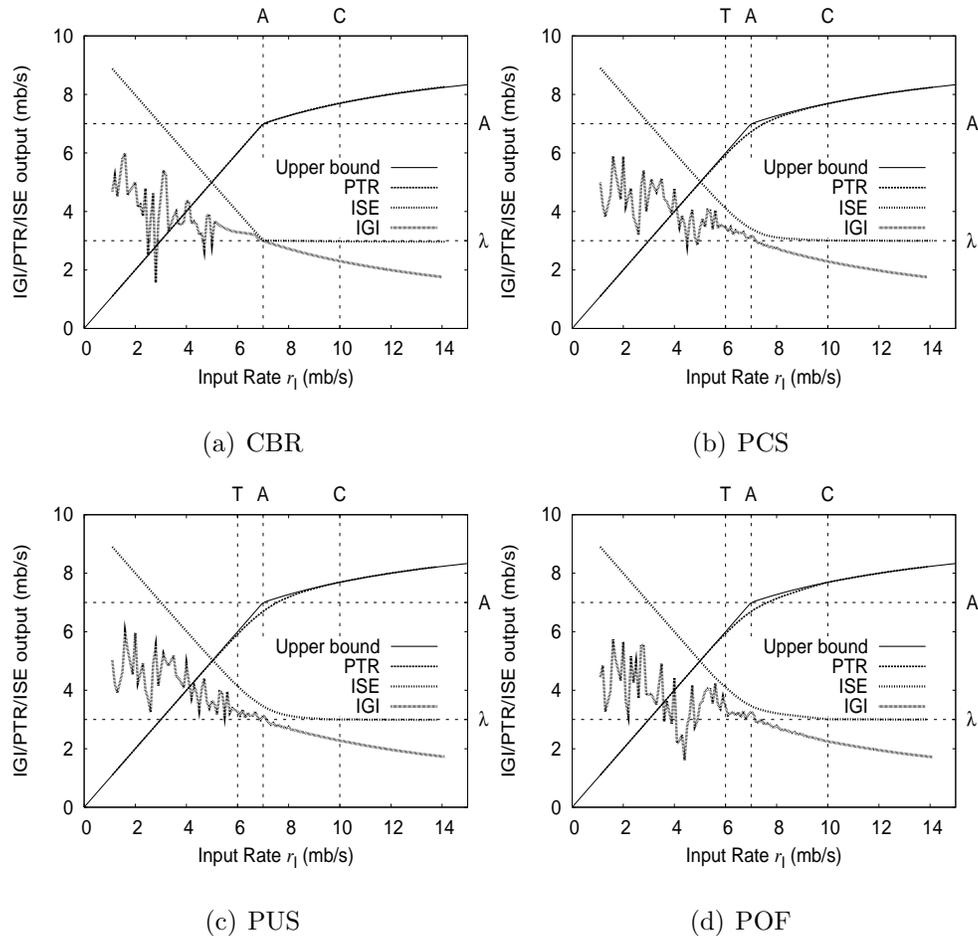
Figure 3.12: Trace driven testing of three estimators: IGI, PTR, and ISE: (a) using CBR, (b) using PCS, (c) using PUS, (d) using POF.

susceptible to cross-traffic interference from non-tight hops, we focus on single-hop analysis in this chapter and skip this issue. A detailed discussion about Spruce is given in the next chapter.

## 3.6 Concluding Remarks

This chapter focused on developing a theoretical understanding of single-hop bandwidth estimation in non-fluid cross-traffic conditions. Our main contri-

butions include a rigorous formulation of all relevant factors in probing-based bandwidth estimation, an analytical methodology featuring intrusion residual analysis, and a thorough discussion of single-hop probing response curves.

Our results clearly delineate the tradeoffs involved in using shorter trains and long trains, which allows choosing the packet-train parameters in a judicious way.

While we identified the *response deviation* as one potential contributing source of measurement errors, there are certainly other important issues related to the performance of measurement techniques such as multi-hop effects, timing errors, and layer-2 effects [33].

In the next chapter, we extend this analysis to multi-hop paths and understand the behavior of current measurement techniques in arbitrary network paths.

# Chapter 4

# Multi-Hop Probing Analysis

## 4.1 Introduction

In this chapter, we extend the asymptotic analysis in chapter 3 to arbitrary network paths and uncover the nature of the measurement bias caused by bursty cross-traffic flows in *multi-hop* network paths. This problem is significantly different from previous single-hop analysis due to the following reasons. First, unlike single-hop measurement, where the input packet-trains have deterministic and equal inter-packet separation formed by the probing source, the input packet-trains at any hop (except the first one) along a multi-link path are output from the previous hop and have random structure. Second and more importantly, the multi-hop probing asymptotics are strongly related to the routing pattern of cross-traffic flows. This is an issue that never arises in a single-hop path and has received little attention in prior investigation. However, as we show in this chapter, it is one of the most significant factors that affect the bandwidth measurement accuracy in multi-hop paths.

To characterize packet-train bandwidth estimation in its most general set-

tings, we derive the probing response curve $\mathcal{Z}$ of a multi-hop path $\mathcal{P}$ assuming arbitrarily routed bursty cross-traffic flows. We compare $\mathcal{Z}$ with its "multi-hop fluid counterpart" $\mathcal{F}$, a response curve obtained when every cross-traffic flow in $\mathcal{P}$ is hypothetically replaced using a CRF flow of the same average intensity and routing pattern. Under an ergodic stationarity assumption for each cross-traffic flow, we show that the real curve $\mathcal{Z}$ is tightly lower bounded by its fluid counterpart $\mathcal{F}$. The curve $\mathcal{Z}$ asymptotically approaches its fluid bound $\mathcal{F}$ in the entire input range as probing packet size or packet-train length increases.

Most of the existing techniques are based on the single-hop fluid response curve $\mathcal{S}$ associated with the bottleneck link in $\mathcal{P}$. Therefore, any deviation of the real curve $\mathcal{Z}$ from the single-hop curve $\mathcal{S}$ can potentially cause measurement bias in bandwidth estimation. Note that the deviation $\mathcal{Z}-\mathcal{S}$ can be decomposed as

$$\mathcal{Z} - \mathcal{S} = (\mathcal{Z} - \mathcal{F}) + (\mathcal{F} - \mathcal{S}). \tag{4.1}$$

The first term $\mathcal{Z} - \mathcal{F}$ is always positive and causes asymptotic underestimation of $A_{\mathcal{P}}$ for most of the existing techniques. This deviation term and its resulting measurement bias are "elastic" in the sense that they can be reduced to an arbitrarily negligible level using packet-trains of sufficient length[1]. For the second deviation term $\mathcal{F} - \mathcal{S}$, we note that both $\mathcal{S}$ and $\mathcal{F}$ are piece-wise linear curves. The first two linear segments in $\mathcal{F}$ associated with *large* input dispersions coincide with $\mathcal{S}$ (i.e., $\mathcal{F} - \mathcal{S} = 0$). The rest of the linear segments in $\mathcal{F}$ associated with *small* input dispersions appear above $\mathcal{S}$ (i.e., $\mathcal{F} - \mathcal{S} > 0$). The amount of deviation and the additional negative measurement bias it causes are dependent on the routing patterns of cross-traffic flows, and are maximized

---

[1]In practice, probing packet-size is limited to 1500 bytes and can not be arbitrarily large.

when every flow traverses only one hop along the path (called one-hop persistent cross-traffic routing). Furthermore, the curve deviation $\mathcal{F} - \mathcal{S}$ is "non-elastic" and stays constant with respect to probing packet size and packet-train length at any given input rate. Therefore, the measurement bias it causes cannot be overcome by adjusting the input packet-train parameters.

Among current measurement techniques, pathload and PTR operate on the input probing range where $\mathcal{F}$ coincides with $\mathcal{S}$, and consequently are only subject to the measurement bias caused by the first deviation term $\mathcal{Z} - \mathcal{F}$. Spruce may use the probing range where $\mathcal{F} - \mathcal{S} > 0$. Hence it is subject to both elastic and non-elastic negative measurement biases.

The rest of the chapter is organized as follows. Section 4.2 derives the multi-hop response curve $\mathcal{F}$ assuming arbitrarily routed CRF cross-traffic flows and examines the deviation term $\mathcal{F} - \mathcal{S}$. In Section 4.3 and 4.4, we derive the real response curve $\mathcal{Z}$ of a multi-hop path and show its relationship to its fluid counterpart $\mathcal{F}$. We provide practical evidence to our theoretical results using simulations, testbed experiments, and real Internet measurements in Section 4.5. We examine the impact of these results on existing techniques in Section 4.6 and conclude in Section 4.7.

## 4.2   Multi-Hop CRF Response Curves

It is important to first thoroughly understand the response curve $\mathcal{F}$ of a network path carrying CRF cross-traffic flows, since as we show later, it is an *approachable* bound of the real response curve $\mathcal{Z}$. Initial investigation of the CRF curves is due to Melandar *et al.* [26] and Dovrolis *et al.* [9]. However, prior work only considers two special cross-traffic routing cases (one-hop persistent routing and

path persistent routing). In this section, we formulate and solve the problem for arbitrary cross-traffic routing patterns, based on which, we discuss several important properties of the CRF response curves that allow us to obtain the path available bandwidth information.

## 4.2.1  Formulation and Solution

We view an $N$-hop network path $\mathcal{P} = (L_1, L_2, \ldots, L_N)$ as a sequence of $N$ interconnected *First-Come First-Served (FCFS) store-and-forward hops*. For each forwarding hop $L_i$ in $\mathcal{P}$, we denote its link capacity by $C_i$, and assume that it has infinite buffer space and a work-conserving queuing discipline. Suppose that $f_1, f_2, \ldots f_M$ are $M$ CRF cross-traffic flows traversing path $\mathcal{P}$. The flow rate of $f_j$ is denoted by $\psi_j$ and the flow rate vector is given by $\Psi = (\psi_1, \psi_2, \ldots, \psi_M)^T$.

**Definition 9** *A* flow aggregation *is a set of flows, represented by a "selection vector"* $\mathbf{p} = (p_1, p_2, \ldots, p_M)^T$, *where* $p_j = 1$ *if flow* $f_j$ *belongs to the flow aggregation and* $p_j = 0$ *if otherwise.*

A single flow $f_j$ is also viewed as a flow aggregation whose selection vector has all but the $j^{th}$ element equal to 0. We use the notation $f_j$ and its selection vector interchangeably. There are several operations between flow aggregations. First, the common flows to aggregations $\mathbf{p}$ and $\mathbf{q}$ form another flow aggregation, whose selection vector is given by $\mathbf{p} \otimes \mathbf{q}$, where the operator $\otimes$ represents "element-wise multiplication". Second, the equality $\mathbf{p} \otimes \mathbf{q} = \mathbf{p}$ implies that aggregation $\mathbf{q}$ contains all flows in aggregation $\mathbf{p}$. Third, the aggregation that contains the flows in $\mathbf{p}$ but not in $\mathbf{q}$ is given by $\mathbf{p} - \mathbf{p} \otimes \mathbf{q}$, where $\otimes$ has higher priority than the subtraction operator. Finally, note that the traffic intensity of a flow aggregation $\mathbf{p}$ can be computed from the inner product $\Psi^T \mathbf{p}$.

We now define several types of flow aggregation frequently used later in this chapter . First, the traversing flow aggregation at link $L_i$, denoted by its selection vector $\mathbf{r}_i$, includes all CRF flows that pass through $L_i$. The $M \times N$ matrix $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is called the "routing matrix" of cross-traffic flows over path $\mathcal{P}$. The selection vector $\mathbf{r}_i$ is also called the routing vector of link $L_i$. For convenience, we define an auxiliary routing vector $\mathbf{r}_0 = \mathbf{0}$.

Before getting to the second type of flow aggregation, we state two routing constraints to simplify discussions. The first constraint requires every flow to have different routing pattern and consequently all row vectors in the routing matrix $\mathbf{R}$ are different among each other. In the case of otherwise, the flows with the same routing pattern should be aggregated into one single flow. The second routing constraint requires every flow to have only one link where it enters the path and also have only one (downstream) link where it exits from the path. In the case of otherwise, the flow is decomposed into several separate flows that meet this routing constraint.

The second type of flow aggregation, denoted by $\mathbf{e}_i$, includes all flows entering the path at link $L_i$, which can be expressed as $\mathbf{e}_i = \mathbf{r}_i - \mathbf{r}_i \otimes \mathbf{r}_{i-1}$ given the second routing constraint stated previously. The third type of flow aggregation, which includes flows that enter the path at link $L_k$ and traverse the downstream link $L_i$, is denoted as $\Gamma_i^k = \mathbf{e}_k \otimes \mathbf{r}_i$, where $k \leq i$.

The cross-traffic intensity at link $L_i$ is denoted by $\lambda_i$. It is assumed to be less than the link capacity $C_i$:

$$\lambda_i = \Psi^T \mathbf{r}_i < C_i, \quad 1 \leq i \leq N. \tag{4.2}$$

We call the two-dimensional vector $\mathbf{h}_i = (C_i, \lambda_i)^T$ the *hop configuration* of $L_i$, the $2 \times N$ matrix $\mathbf{H} = (\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_N)$ the *path configuration* of $\mathcal{P}$. The

two row vectors in $\mathbf{H}$ are referred to as capacity vector and intensity vector, denoted by $\mathbf{C}$ and $\mathbf{\Lambda}$ respectively. The hop available bandwidth of $L_i$ is given by $A_i = C_i - \lambda_i$. We assume that every hop has different available bandwidth, and consequently that the tight link is unique. Sometimes, we need to refer to the second minimum hop available bandwidth and the associated link, which we denote as $A_{b2} = C_{b2} - \lambda_{b2}$ and $L_{b2}$ respectively. That is

$$b2 = \arg \min_{1 \leq i \leq N, i \neq b} (C_i - \lambda_i), \tag{4.3}$$

where $b$ is the index of the tight hop.

When a packet-train of input dispersion (i.e., inter-packet spacing) $g_I$ and packet size $s$ is used to probe path $\mathcal{P}$, we are interested in computing the output dispersion of the packet train and examining its relation to $g_I$. Such a relation is called the gap response curve of path $\mathcal{P}$. It is easy to verify that under the CRF condition, for any cross-traffic routing matrix $\mathbf{R}$, the response curve does not depend on the packet-train length $n$. Hence, we only consider the case of packet-pair probing. We denote the output dispersion at link $L_i$ as $\gamma_i(g_I, s)$ or $\gamma_i$ for short, and again for notational convenience we let $\gamma_0 = g_I$. Note that $\gamma_N(g_I, s)$ is a more elaborate version of the notation $\mathcal{F}$ we have used previously.

Based on the above formulation, the gap response curve of path $\mathcal{P}$ have a recursive representation given in the following theorem.

**Theorem 10** *When a packet-pair with input dispersion $g_I$ and packet size $s$ is used to probe an N-hop CRF path $\mathcal{P}$ with routing matrix $\mathbf{R}$ and flow rate vector $\Psi$, the output dispersion at link $L_i$ can be recursively expressed as follows*

$$\gamma_i = \begin{cases} g_I & i = 0 \\ \max\left(\gamma_{i-1}, \dfrac{s + \Omega_i}{C_i}\right) & i > 0 \end{cases}, \tag{4.4}$$

where $\Omega_i$ is given by

$$\Omega_i = \sum_{k=1}^{i} \left[ \gamma_{k-1} \Psi^T \Gamma_i^k \right]. \tag{4.5}$$

**Proof:** Note that the term $\Omega_i$ represents the volume of fluid cross-traffic buffered between the packet-pair in the outgoing queue of link $L_i$. As an analogy, we can view packet-pair as a bus, cross-traffic as passengers, and routers as bus stations. Then, $\Omega_i$ is the cross-traffic that is "picked up" by the packet-pair at link $L_i$ as well as all the upstream links of $L_i$; and they will traverse over link $L_i$ due to the flows' routing decision.

We now prove (4.4). Assuming that the first packet arrives at link $L_i$ at time instance $a_1$. It gets immediate transmission service and departs at $a_1 + s/C_i$. The second packet arrives at $a_1 + \gamma_{i-1}$. The server of $L_i$ needs to transmit $s + \Omega_i$ amount of data before it can serve the second packet. If this is done before time instance $a_1 + \gamma_{i-1}$, the second packet also gets immediate service and $\gamma_i = \gamma_{i-1}$. Otherwise, the sever undergoes a busy period between the departure of the two packets, meaning that $\gamma_i = (s + \Omega_i)/C_i$. Therefore, we have

$$\gamma_i = \max\left( \gamma_{i-1}, \frac{s + \Omega_i}{C_i} \right). \tag{4.6}$$

This completes the proof of the theorem.  ∎

As a quick sanity check, we verify the compatibility between Theorem 10 and the special "one-hop persistent" routing case. In one-hop persistent routing [10], every flow that enters the path at link $L_i$ will exit the path at link $L_{i+1}$. The element-wise multiplication of any two different routing vectors gives a zero vector. That is, $\mathbf{r}_i \otimes \mathbf{r}_k = \mathbf{0}$ for $i \neq k$. Hence, we have

$$\Gamma_i^k = \mathbf{r}_k \otimes \mathbf{r}_i - \mathbf{r}_k \otimes \mathbf{r}_{k-1} \otimes \mathbf{r}_i = \begin{cases} \mathbf{0} & i \neq k \\ \mathbf{r}_i & i = k \end{cases}. \tag{4.7}$$

Therefore, (4.5) can be simplified as

$$\Omega_i = \gamma_{i-1}\Psi^T \mathbf{r}_i = \gamma_{i-1}\lambda_i, \tag{4.8}$$

which agrees with previous results [26] [9].

## 4.2.2   Properties of CRF Response Curves

Theorem 10 leads to several important properties of the CRF response curve, which we discuss next. Note that none of these results for arbitrary cross-traffic routing have been confirmed in previous work.

**Property 1** *The output dispersion $\gamma_N(g_I, s)$ is a continuous piece-wise linear function of the input dispersion $g_I$ in the input dispersion range $(0, \infty)$.*

**Proof:**   We apply mathematical induction to $i$. When $i = 0$, according to the first formula in (4.4), $\gamma_0 = g_I$ is a continuous linear function of $g_I$. Assuming for any $0 \le i < N$, $\gamma_i$ is a continuous piece-wise linear function of $g_I$, we show that $\gamma_N$ is also a continuous piece-wise linear function of $g_I$. From (4.5), we know that $\Omega_N$ is a linear combination of $\gamma_i$, where $1 \le i < N$. Therefore, $\Omega_i$ is a continuous piece-wise linear function of $g_I$. Combining this result with the second part of (4.4) and the induction hypothesis which states the piece-wise linearity of $\gamma_{N-1}$, the desired property follows for $\gamma_N$.   ■

Let $0 = \alpha_{K+1} < \alpha_K < \ldots < \alpha_1 < \alpha_0 = \infty$ be the input dispersion turning points that split the gap response curve to $K + 1$ linear segments. To understand what the linear segments and the turning points are associated with, we introduce a concept called "*congestible hop set.*"

**Definition 10** *When a packet-pair with input dispersion $g_I$ and packet size $s$ is injected into the CRF path $\mathcal{P}$, the set of hops at which the dispersion of the*

*traversing packet-pair gets expanded is called the* congestible hop set *of path $\mathcal{P}$ at input rate $s/g_I$. It is easy to verify that the congestible hop set is decided by the ratio of $s$ to $g_I$ and not their individual values. We denote the congestible hop set as $\mathcal{H}(\mathcal{P}, s/g_I)$:*

$$\mathcal{H}(\mathcal{P}, s/g_I) = \{L_i : L_i \in \mathcal{P} \wedge \gamma_{i-1}(g_I, s) < \gamma_i(g_I, s)\}. \tag{4.9}$$

**Lemma 8** *A necessary and sufficient condition for two input dispersions $g_1$ and $g_2$ to fall into the same linear segment of the gap response curve is*

$$\mathcal{H}(\mathcal{P}, s/g_1) = \mathcal{H}(\mathcal{P}, s/g_2). \tag{4.10}$$

Among the turning points, the first two $\alpha_1$ and $\alpha_2$ are of major interests. Among the linear segments, the first segment in the input dispersion range $(\alpha_1, \infty)$ and the second segment in the range $(\alpha_2, \alpha_1)$ are most important for bandwidth estimation.

**Property 2** *The first turning point $\alpha_1$ corresponds to the path available bandwidth in the sense that $A_{\mathcal{P}} = s/\alpha_1$. The first linear segment in the input dispersion range $(\alpha_1 = s/A_{\mathcal{P}}, \infty)$ has slope 1 and intercept 0. The second linear segment in the input dispersion range $(\alpha_2, \alpha_1)$ has slope $\lambda_b/C_b$ and intercept $s/C_b$, where $b$ is the index of the tight link:*

$$\gamma_N(g_I, s) = \begin{cases} g_I & \alpha_1 \leq g_I \leq \infty \\ \dfrac{g_I \lambda_b + s}{C_b} & \alpha_2 \leq g_I \leq \alpha_1 \end{cases}. \tag{4.11}$$

*These facts are irrespective of the routing matrix.*

**Proof:** To prove the first part in (4.11), we apply mathematical induction to $i$ to show that there is no link at which the packet-pair gets expanded when $g_I \geq s/A_{\mathcal{P}}$. For $i = 0$, $\gamma_0 = g_I$ due to the first part of (4.4).

Suppose that for $0 \leq i < N$, $\gamma_i = g_I$, we show that $\gamma_N = g_I$. Combining induction hypothesis and (4.5), we get

$$\Omega_N = \sum_{k=1}^{N} \left[ \gamma_{k-1} \Psi^T \Gamma_i^k \right] = g_I \Psi^T \sum_{k=1}^{N} \Gamma_i^k$$
$$= g_I \Psi^T \mathbf{r}_N = g_I \lambda_N. \tag{4.12}$$

Further recalling that

$$\frac{s}{g_I} \leq A_{\mathcal{P}} \leq A_N = C_N - \lambda_N, \tag{4.13}$$

we have $(s + g_I \lambda_N)/C_N \leq g_I = \gamma_{N-1}$. Combining (4.4), we have $\gamma_N = g_I$.

When $\alpha_1 < g_I \leq \alpha_2$, there is only the tight link $L_b$ that expands the packet-pair dispersion. Due to the same derivation as in (4.12), $\Omega_b = g_I \lambda_b$. Combining the second part in (4.4) and the fact that $s/g_I > A_p = C_b - \lambda_b$, we get

$$\gamma_b = \max\left( g_I, \frac{s + g_I \lambda_b}{C_b} \right) = \frac{s + g_I \lambda_b}{C_b}. \tag{4.14}$$

Finally, notice that $\gamma_N = \gamma_b$ due to the fact that $L_b$ is the only link that expands the packet-pair. ∎

It helps to find the expression for the turning point $\alpha_2$, so that we can identify the exact range for the second linear segment. However, unlike $\alpha_1$, the turning point $\alpha_2$ is dependent on the routing matrix. In fact, all other turning points are dependent on the routing matrix and can not be computed based on the path configuration matrix alone. Therefore, we only provide a bound for $\alpha_2$.

**Property 3** *For any routing matrix, the term $s/\alpha_2$ is no less than $A_{b2}$, which is the second minimum hop available bandwidth of path $\mathcal{P}$.*

The slopes and intercepts for all but the first two linear segments are related to the routing matrix. We skip the derivation of their expressions, but instead provide both a lower bound and an upper bound for the entire response curve.

**Property 4** *For a given path configuration matrix, the gap response curve associated with any routing matrix is lower bounded by the following piece-wise linear function*

$$
\mathcal{S}(g_I, s) = \begin{cases} g_I & g_I > \dfrac{s}{A_{\mathcal{P}}} \\ \dfrac{s + g_I \lambda_b}{C_b} & 0 < g_I < \dfrac{s}{A_{\mathcal{P}}} \end{cases}.
\tag{4.15}
$$

*It is upper bounded by the gap response curve associated with one-hop persistent routing.*

**Proof:** The lower bound is obvious, so we only prove the upper bound. We apply mathematical induction to show that the output dispersion $\gamma_N^p(g_I, s)$ associated with one-hop persistent routing, is no less than the output dispersion $\gamma_N(g_I, s)$ associated with any other cross-traffic routing patterns. For $i = 0$, $\gamma_0^p = \gamma_0 = g_I$. Assuming for all $i < N$, we have $\gamma_i^p \geq \gamma_i$. We next show that $\gamma_N^p \geq \gamma_N$. First, we establish the fact that $\Omega_N \leq \gamma_{N-1}\lambda_N$ as follows

$$
\Omega_N = \sum_{k=1}^{N} \left[ \gamma_{k-1} \Psi^T \Gamma_i^k \right] \leq \gamma_{N-1} \sum_{k=1}^{N} \left[ \Psi^T \Gamma_i^k \right]
$$
$$
= \gamma_{N-1} \Psi^T \mathbf{r}_N = \gamma_{N-1}\lambda_N.
\tag{4.16}
$$

Then we have $\gamma_N^p \geq \gamma_N$ due to the following

$$
\gamma_N(g_I, s) = \max\left( \gamma_{N-1}, \frac{s + \Omega_N}{C_N} \right)
$$
$$
\leq \max\left( \gamma_{N-1}, \frac{s + \lambda_N \gamma_{N-1}}{C_N} \right)
$$
$$
\leq \max\left( \gamma_{N-1}^p, \frac{s + \lambda_N \gamma_{N-1}^p}{C_N} \right) = \gamma_N^p,
\tag{4.17}
$$

where the second inequality follows from the induction hypothesis. ∎

Recall that the lower bound $\mathcal{S}(g_I, s)$ is the single-hop response curve of the bottleneck link $L_b$, denoted in short by $\mathcal{S}$ in the introduction. We now make several observations regarding the deviation of $\gamma_N(g_I, s)$ (i.e., $\mathcal{F}$) from $\mathcal{S}(g_I, s)$. Combing (4.11) and (4.15), we see that $\gamma_N(g_I, s) - \mathcal{S}(g_I, s) = 0$ for $g_I \geq \alpha_2$. That is, the first two linear segments on $\mathcal{F}$ coincide with $\mathcal{S}$. For $g_I < \alpha_2$, Property 4 implies that the deviation $\gamma_N(g_I, s) - \mathcal{S}(g_I, s)$ is positive. The exact value depends on cross-traffic routing and it is maximized in one-hop persistent routing given a fixed path configuration matrix.

Note that there are three pieces of path information that we can extract from the gap response curve $\mathcal{F}$ without resorting to the routing matrix, which is usually not available in practice. By locating the first turning point $\alpha_1$, we can compute the path available bandwidth. From the second linear segment, we can obtain the tight link capacity and cross-traffic intensity (and consequently, the bottleneck link utilization) information. Other parts of the response curve $\mathcal{F}$ are less readily usable due to their dependence on the routing matrix.

To extract bandwidth information from the output dispersion $\gamma_N$, it is often more helpful to look at the *rate* response curve, i.e., the functional relation between the output rate $r_O = s/\gamma_N$ and the input rate $r_I = s/g_I$. However, since this relation is not linear, we adopt a transformed version first proposed by Melander *et al.* [28], which depicts the relation between the ratio $r_I/r_O$ and $r_I$. Denoting this rate response curve by $\tilde{F}(r_I)$, we have

$$\tilde{F}(r_I) = \frac{r_I}{r_O} = \frac{\gamma_N(g_I, s)}{g_I}. \tag{4.18}$$

Suppose that the $j^{th}$ linear segment in the gap response curve is expressed as $\gamma_N(g_I, s) = \ell_j g_I + \kappa_j$, then in the input rate range $(s/\alpha_{j-1}, s/\alpha_j)$, the rate

response curve $\tilde{F}(r_I)$ can be expressed as

$$\tilde{F}(r_I) = \frac{\gamma_N(g_I, s)}{g_I} = \ell_j + \frac{\kappa_j}{g_I} = \ell_j + \frac{\kappa_j}{s} r_I. \tag{4.19}$$

This shows that $\tilde{F}(r_I)$ is also piece-wise linear. It is easy to see that the first turning point in the rate curve is $s/\alpha_1 = A_p$ and that the rate curve in the input rate range $(0, s/\alpha_2)$ can be expressed as follows:

$$\tilde{F}(r_I) = \begin{cases} 1 & r_I \leq A_{\mathcal{P}} \\ \dfrac{\lambda_b + r_I}{C_b} & \dfrac{s}{\alpha_2} \geq r_I \geq A_{\mathcal{P}} \end{cases}. \tag{4.20}$$

Finally, we state the following important property for the rate response curve $\tilde{F}(r_I)$.

**Property 5** *For any given input rate $r_I \in (0, \infty)$, the rate response $\tilde{F}(r_I)$ does not depend on probing packet size $s$.*

**Proof:** First, by applying induction on $i$, we can prove that for any given $r_I$, $\gamma_N(g_I, s)$ is proportional to $s$. We skip the details for this step. Further notice that $g_I = s/r_I$ is also proportional to $s$. Combining theses facts and (4.18), the property follows. ∎

### 4.2.3 Examples and Discussions

We use a simple example to illustrate the response curve properties. Suppose that we have a 3-hop path with equal capacity 10mb/s. We consider two routing matrices and flow rate settings that lead to the same path configuration.

In the first setting, the flow rate vector $\Psi$ is $(4, 7, 8)^T$ and the routing pattern

is one-hop persistent. That is,

$$\Lambda^T = \mathbf{R}^T \Psi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^T \times \begin{pmatrix} 4 \\ 7 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 8 \end{pmatrix}. \tag{4.21}$$

In the second setting, the flow rate vector $\Psi$ is $(4, 3, 1)^T$ and the routing pattern is path persistent. That is,

$$\Lambda^T = \mathbf{R}^T \Psi = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^T \times \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 8 \end{pmatrix}. \tag{4.22}$$

The probing packet size $s$ is 1500 bytes. The CRF gap response curves for both routing patterns are plotted in Fig. 4.1(a). In this example, both curves have 4 linear segments separated by turning points $\alpha_1 = 6$ms, $\alpha_2 = 4$ms, and $\alpha_3 = 2$ms. It is provable that when links are arranged in the decreasing order of their hop available bandwidth, every turning point corresponds to a hop available bandwidth regardless of the cross-traffic routing. Note that in Fig. 4.1(a), the curve for path-persistent routing appears below the one for one-hop persistent routing. The lower bound $\mathcal{S}$ identified in Property 4 is also shown in the figure. This lower bound is the gap response curve of the single-hop path comprising only the tight link of $\mathcal{P}$.

The rate response curves for the two examples are given in Fig. 4.1(b), where the three turning points are 2mb/s, 3mb/s, and 6mb/s respectively. Due to the transformation we adopted, the rate curve for one-hop persistent routing still remains as an upper bound for the rate curves associated with the other routing patterns. From Fig.4.1(b), we also see that, similar to the gap curves, the two multi-hop rate response curves and their lower bound $\tilde{S}(r_I)$ (i.e., the

(a) gap response curve  (b) rate response curve
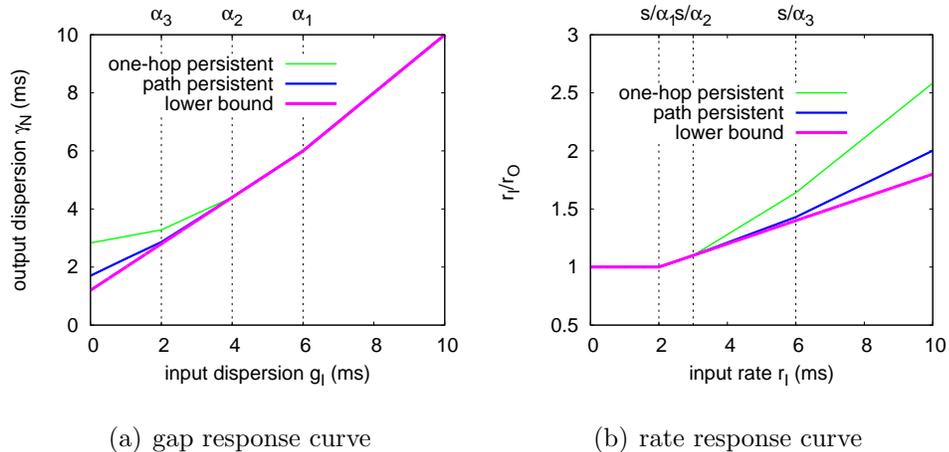
Figure 4.1: An example of multi-hop response curves.

transformed rate version of $\mathcal{S}(g_I, s)$) share the same first and second linear segments.

We conclude this section by discussing several major challenges in extending the response curve analysis to a multi-hop path carrying *bursty* cross-traffic flows. First, notice that with bursty cross-traffic, even when the input dispersion and packet-train parameters remain constant, the output dispersion is no longer deterministic as it is under the CRF condition. Rather, it becomes a random variable. Accordingly, we define the gap response curve $\mathcal{Z}$ to be the functional relation between the input dispersion and the statistical mean of the output dispersion random variable. Second, unlike in the CRF case, where both packet-train length $n$ and probing packet size $s$ have no impact on the rate response curve $\tilde{F}(r_I)$, the response curves in bursty cross-traffic are strongly related to these two packet-train parameters. Finally, a full characterization of a CRF flow only requires one parameter – its arrival rate, while a full characterization of a bursty flow requires several stochastic processes. In what follows, we address these problems and extend our analysis to multi-hop paths with bursty cross-

Table 4.1: Random Process Notations

| | |
|---|---|
| $\{V_i(\mathbf{p}, t)\}$ | The cumulative arrival process of $\mathbf{p}$ at link $L_i$ |
| $\{Y_\delta^i(\mathbf{p}, t)\}$ | The cross-traffic intensity process $\mathbf{p}$ at $L_i$ |
| $\{W_i(\mathbf{p}, t)\}$ | The hop workload process at $L_i$ w.r.t. $\mathbf{p}$ |
| $\{D_\delta^i(\mathbf{p}, t)\}$ | The workload-difference process at $L_i$ w.r.t. $\mathbf{p}$ |
| $\{U_i(\mathbf{p}, t)\}$ | The hop utilization process at $L_i$ w.r.t. $\mathbf{p}$ |
| $\{B_\delta^i(\mathbf{p}, t)\}$ | The available bandwidth process at $L_i$ w.r.t. $\mathbf{p}$ |

traffic.

## 4.3 Basics in Multi-Hop Analysis

In this section, we present a stochastic formulation for the multi-hop bandwidth measurement problem and derive a recursive expression for the output dispersion random variable. This expression is a fundamental result that the asymptotic analysis in Section 4.4 is based upon.

### 4.3.1 Formulation

We keep most of the formulation for CRF path, with some of the terms having different meanings explained in the while. Since cross-traffic flows now become bursty flows of data packets, we adopt the definitions of several random processes (Definition 1-6) in chapter 3 to characterize them. However, these definitions need to be refined so that they are with respect to a particular hop and a particular flow aggregation. In what follows, we refine the two basic random processes, which the definitions of the other processes are based upon. The notations for all the six random processes are given in Table 4.1.

**Definition 11** *The cumulative traffic arrival of flow aggregation* **p** *at link* $L_i$, *denoted as* $\{V_i(\mathbf{p}, t), 0 \leq t < \infty\}$ *is a random process counting the total volume of data (in bits) received by hop* $L_i$ *from flow aggregation* **p** *up to time instance* $t$.

**Definition 12** *Hop workload process of* $L_i$ *with respect to flow aggregation* **p**, *denoted as* $\{W_i(\mathbf{p}, t), 0 \leq t < \infty\}$ *indicates the sum at time instance t of service times of all packets in the queue and the remaining service time of the packet in service, under such a* hypothetical *situation that the server of* $L_i$ *only transmits packets from flow aggregation* **p**, *while silently drops all packets not from* **p**.

We next make several modeling assumptions on cross-traffic flows. First, we assume all flows have stationary arrivals.

**Assumption 3** *For any cross-traffic flow* $f_j$ *that enters the path from link* $L_i$, *the cumulative traffic arrival process* $\{V_i(f_j, t)\}$ *has ergodic stationary increments. That is, for any* $\delta > 0$, *the* $\delta$-*interval traffic intensity process* $\{Y_\delta^i(f_j, t)\}$ *is a mean-square ergodic process with time-invariant distribution and ensemble mean* $\psi_j$.

We explain this assumption in more details. First, the stationary increment assumption implies that the increment process of $\{V_i(f_j, t)\}$ for any given time interval $\delta$, namely $\{V_i(f_j, t + \delta) - V_i(f_j, t) = \delta Y_\delta^i(f_j, t)\}$, has time-invariant distribution. This further implies that the $\delta$-interval traffic intensity process $\{Y_\delta^i(f_j, t)\}$ also has time-invariant distribution. Second, the mean-square ergodicity can be mathematically expressed as follows

$$\lim_{\tau \to \infty} E\left[\left(\frac{1}{\tau}\int_0^\tau Y_\delta^i(f_j, u)du - \psi_j\right)^2\right] = 0. \tag{4.23}$$

It is easy to prove that, an equivalent expression of (4.23) is given by

$$\lim_{\delta \to \infty} E\left[\left(Y_\delta^i(f_j, t) - \psi_j\right)^2\right] = 0 \quad \forall t > 0. \tag{4.24}$$

This implies that, for any $t$, as the observation interval $\delta$ increases, the random variable $Y_\delta^i(f_j, t)$ converges to $\psi_j$ in mean-square sense. In other words, the variance of $Y_\delta^i(f_j, t)$ decays to 0.

Our next assumption makes sure that every hop has sufficient capacity to accommodate the traversing flows and consequently that the queue at each link remains stable.

**Assumption 4** *For any link $L_i$, the traversing flows have an aggregated traffic intensity less than the link capacity $C_i$. That is, $\lambda_i = \Psi^T \mathbf{r}_i < C_i$.*

As a consequence of the two assumptions, the ergodic stationary property also holds for any flow aggregations at their entering link, as stated formally in the following.

**Corollary 5** *For any flow aggregation $\mathbf{p}$ that enters the path at link $L_i$, i.e., $\mathbf{p} \otimes \mathbf{e}_i = \mathbf{p}$, the process $\{V_i(\mathbf{p}, t)\}$ has ergodic stationary increments. That is, the traffic intensity process $\{Y_\delta^i(\mathbf{p}, t)\}$ has time-invariant distribution with ensemble mean $\Psi^T \mathbf{p}$ and the mean-square convergence expressed by the following*

$$\lim_{\delta \to \infty} E\left[\left(Y_\delta^i(\mathbf{p}, t) - \Psi^T \mathbf{p}\right)^2\right] = 0 \quad \forall t > 0. \tag{4.25}$$

We point out that the correlation among flows in any aggregation does not affect the validity of Corollary 5. This is because when observation interval $\delta$ is sufficiently large, the intensity process $Y_\delta^i(f_j, t)$ for any flow has negligible variance. Consequently, the correlation among flows also becomes negligible. It

is also important to notice that flow correlation does not concern any results in this chapter, whose focus is on the asymptotic situations where the correlation between any flows becomes negligible.

Due to Szczotka [39] [40], the workload process $\{W_i(\mathbf{p}, t)\}$ will "inherit" the ergodic stationarity property from the traffic arrival process $\{V_i(\mathbf{p}, t)\}$. This property is further carried over to the $\delta$-interval workload-difference process $\{D_\delta^i(\mathbf{p}, t)\}$ and the available bandwidth process $\{B_\delta^i(\mathbf{p}, t)\}$, whose ensemble means are 0 and $A_i(\mathbf{p}) = C_i - \Psi^T \mathbf{p}$ respectively[2]. Further, the ergodicity property leads to the following result.

**Lemma 9** *For any flow aggregation* $\mathbf{p}$ *that enter the path at link* $L_i$, *the random variable* $B_\delta^i(\mathbf{p}, t)$ *converges in mean-square sense to* $C_i - \Psi^T \mathbf{p}$ *as* $\delta \to \infty$.

$$\lim_{\delta \to \infty} E\left[\left(B_\delta^i(\mathbf{p}, t) - \left(C_i - \Psi^T \mathbf{p}\right)\right)^2\right] = 0 \quad \forall t > 0. \tag{4.26}$$

On the other hand, notice that unlike $\{Y_\delta^i(\mathbf{p}, t)\}$ and $\{B_\delta^i(\mathbf{p}, t)\}$, the workload difference process $\{D_\delta^i(\mathbf{p}, t)\}$ is not a moving average process by nature. Consequently, the mean-square ergodicity of $\{D_\delta^i(\mathbf{p}, t)\}$ does not cause the variance of $D_\delta^i(\mathbf{p}, t)$ to decay with respect to the increase of $\delta$. Instead, we have the following lemma.

**Lemma 10** *For any* $t$, *the variance of the random variable* $D_\delta^i(\mathbf{p}, t)$ *converges to* $2Var[W_i(\mathbf{p}, t)]$ *as* $\delta$ *increases.*

$$\lim_{\delta \to \infty} E\left[\left(D_\delta^i(\mathbf{p}, t) - 0\right)^2\right] = 2Var\left[W_i(\mathbf{p}, t)\right]. \tag{4.27}$$

---

[2] The available bandwidth of $L_i$ is $A_i = A_i(\mathbf{r}_i) = C_i - \Psi^T \mathbf{r}_i$, and not $A_i(\mathbf{p})$.

**Proof:** Due to the definition of workload-difference process and the stationarity of workload process, we have

$$E[(D_\delta^i(t))^2] = E[(W_i(t+\delta) - W_i(t))^2]$$

$$= E[(W_i(t+\delta))^2 - 2W_i(t)W_i(t+\delta) + (W_i(t))^2]$$

$$= 2\left(E[(W_i(t))^2] - E[W_i(t)W_i(t+\delta)]\right). \tag{4.28}$$

Notice that due to the mean-square ergodicity of workload process, the random variable $W_i(t+\delta)$ are asymptotically uncorrelated to $W_i(t)$ as $\delta$ increase. That is

$$\lim_{\delta\to\infty} E[(W_i(t) - E[W_i(t)])(W_i(t+\delta) - E[W_i(t+\delta)])] = 0. \tag{4.29}$$

As a consequence of (4.29), we get

$$\lim_{\delta\to\infty} E[W_i(t)W_i(t+\delta)]$$

$$= \lim_{\delta\to\infty} E[W_i(t)]E[W_i(t+\delta)] = E^2[W_i(t)]. \tag{4.30}$$

Taking the limit of (4.28) and combining (4.30), we get (4.27). ∎

In our later analysis, not only we need to know the asymptotic variance of $Y_\delta^i(\mathbf{p},t)$, $D_\delta^i(\mathbf{p},t)$ and $B_\delta^i(\mathbf{p},t)$ when $\delta$ approaches infinity, but also we often rely on such a condition that their variance is uniformly bounded by some constant for any $\delta$. We now justify this condition from a practical standpoint. First note that cross-traffic arrival rate is bounded by the capacities of incoming links at a hop. Suppose that the sum of all incoming link capacities at hop $L_i$ is $\mathcal{C}_i$, then $Y_\delta^i(\mathbf{p},t)$ is distributed in a finite interval $[0,\mathcal{C}_i]$ and its variance is uniform bounded by the constant $\mathcal{C}_i^2$ for any observation interval $\delta$. Similarly, the variance of $B_\delta^i(\mathbf{p},t)$ is uniformly bounded by the constant $C_i^2$. The variance of $D_\delta^i(\mathbf{p},t)$ is uniformly bounded by the constant $4Var[W_i(\mathbf{p},t)]$ for any $\delta$, which directly follows from the definition of $D_\delta^i(\mathbf{p},t)$.

We also point out that some of the notations introduced to formulate CRF path now have different meanings. The rate of the bursty cross-traffic flow $f_j$, denoted by $\psi_j$, is the ensemble mean of the traffic intensity process $\{Y_\delta^i(f_j, t)\}$, which is also the *long-term average* arrival rate of $f_j$ at any link it traverses. The term $\lambda_i = \Psi^T \mathbf{r}_i$ becomes the long-term average arrival rate of the aggregated cross-traffic at link $L_i$. The term $A_i = C_i - \lambda_i$ is the long-term average hop available bandwidth at link $L_i$. Again recall that we explicitly target the measurement of the long-term averages of available bandwidth and/or cross-traffic intensity, instead of those metrics in certain time intervals.

We next introduce several notations to describe packet-train probing. As in [23], we use the quadruple $\langle \{T_m\}, g_I, s, n \rangle$ to denote an infinite packet-train series with input inter-packet dispersion $g_I$, packet size $s$, and packet-train length $n$. This series is driven by a point process $\Lambda(t) = \max\{m \geq 0 : T_m \leq t\}$ to probe path $\mathcal{P}$. Let $d_1(m, i)$ and $d_n(m, i)$ be the departure time instances from link $L_i$ of the first and last probing packets in the $m^{th}$ packet-train. We define the *sampling interval* of the packet-train as the total spacing $\Delta = d_n(m, i) - d_1(m, i)$, and the *output dispersion* as the average spacing $G = \Delta/(n-1)$ of the packet-train. Both $\Delta$ and $G$ are random variables, whose distributions depend on several factors such as the input dispersion $g_I$, the packet-train parameters $s$ and $n$, the packet-train index $m$ in the probing series, the hop $L_i$ that the output dispersion $G$ is associated with. Therefore, a full version of $G$ is written as $G_i(g_I, s, n, m)$. However, for notation brevity, we often omit the parameters that are implicitly known or have little relevance to the topic under discussion.

We now formally state the questions we address in this chapter. Note that a realization of the stochastic process $\{G_N(g_I, s, n, m), 1 \leq m < \infty\}$ is just a packet-train probing experiment. We examine the sample-path time-average

and its relationship to $g_I$ when keeping $s$ and $n$ constant. This relationship, previously denoted by $\mathcal{Z}$, is called the "gap response curve" of path $\mathcal{P}$. In particular, we compare the gap response curve of $\mathcal{P}$ with that of the *CRF counterpart* of $\mathcal{P}$ and prove that the former is lower-bounded by the later.

**Definition 13** *Suppose that path $\mathcal{P}$ has routing matrix $\mathbf{R}$ and flow rate vector $\Psi$ and that path $\overline{\mathcal{P}}$ has routing matrix $\overline{\mathbf{R}}$ and flow rate vector $\overline{\Psi}$. $\overline{\mathcal{P}}$ is called the CRF counterpart of $\mathcal{P}$ if 1) all cross-traffic flows traversing $\overline{\mathcal{P}}$ are constant-rate fluid; 2) the two paths $\overline{\mathcal{P}}$ and $\mathcal{P}$ have the same configuration matrix; and 3) there exists a row-exchange matrix $T$, such that $T\mathbf{R} = \overline{\mathbf{R}}$ and $T\Psi = \overline{\Psi}$.*

From this definition, we see that for every flow $f_j$ in $\mathcal{P}$, there is a corresponding CRF flow $f_{j'}$ in the CRF counterpart of $\mathcal{P}$ such that $f_{j'}$ have the same average intensity and routing pattern as those of $f_j$. Note that the third condition in Definition 13 is made to allow the two flows have different indices, i.e., to allow $j \neq j'$.

A second focus of this chapter is to study the impact of packet-train parameters $s$ and $n$ on the response curves. That is, for any given input rate $r_I$ and other parameters fixed, we examine the convergence properties of the two processes $\{G_N(s/r_I, s, n, m), 0 < s < \infty\}$ and $\{G_N(s/r_I, s, n, m), 2 \leq n < \infty\}$.

Next, we first obtain a basic result for the process $\{G_N(g_I, s, n, m), 1 \leq m < \infty\}$, called "output dispersion process".

## 4.3.2 Analysis of Output Dispersion Process

We keep input packet-train parameters $g_I$, $s$, and $n$ constant and denote for short the output dispersion process at link $L_i$ by $\{G_i(m), 1 \leq m < \infty\}$. To study

the properties of this process, we first derive the expressions for the random variable $G_i(m)$.

**Lemma 11** *Let $G_0(m) = g_I$, the random variable $G_i(m)$ has a recursive expression given in the following*

$$
\begin{aligned}
G_i(m) &= \sum_{k=1}^{i} \frac{Y_{\Delta_{k-1}}^{k}(\Gamma_i^k, t_k) \times G_{k-1}(m)}{C_i} + \frac{s}{C_i} + \frac{\tilde{I}_i(m)}{n-1} \\
&= G_{i-1}(m) + \frac{D_{\Delta_{i-1}}^{i}(\mathbf{e}_i, t_i)}{n-1} + \frac{R_i(m)}{n-1},
\end{aligned}
\tag{4.31}
$$

*where $\Delta_{k-1} = (n-1) \times G_{k-1}(m)$ is the sampling interval of the input packet train at link $L_k$, $t_k$ is the arrival time of the packet-train at $L_k$. The term $R_i(m)$ represents the extra queuing delay (besides $W_i(\mathbf{e}_i, t_i + \Delta_{i-1})$) experienced at $L_i$ by the last probing packet in the train. The term $\tilde{I}_i(m)$ is the hop idle time of $L_i$ during the time interval $[t_i, t_i + \Delta_{i-1}]$.*

**Proof:** For the first equality in (4.31), note that the following term (which is a random variable)

$$
Y_{\Delta_{k-1}}^{k}(\Gamma_i^k, t_k) \times (n-1) \times G_{k-1}(m)
\tag{4.32}
$$

denoted separately by $\Omega_i^k$, is the amount of cross-traffic the packet-train picked up at link $L_k$ that will traverse link $L_i$. The random variable

$$
\Omega_i = \sum_{k=1}^{i} \Omega_i^k
\tag{4.33}
$$

gives the total amount of cross-traffic that $L_i$ has to transmit between the departures of the first and last packets in the packet-train. During that time interval, the server also needs to transmit $n-1$ probing packets, which takes $(n-1)s/C_i$ time units, and idle for $\tilde{I}_i(m)$ time units. Therefore, we have

$$
(n-1) \times G_i(m) = \frac{\Omega_i + (n-1)s}{C_i} + \tilde{I}_i(m).
\tag{4.34}
$$

Dividing by $n-1$ at both sides of (4.34), we get the first equality in (4.31).

For the second equality in (4.31), note that the term $R_i(m)$ is the intrusion residual[3] experienced by the last packet in the probing train. It is the amount of extra queuing delay caused by all but the last probing packets in the packet-train and the cross-traffic packets picked up by packet-train at the upstream links of $L_i$. Let $q_1$ and $q_n$ be the queuing delays experienced by the first and last packet in the train, we have

$$q_1 \quad = \quad W_i(\mathbf{e}_i, t_i) \tag{4.35}$$

$$q_n \quad = \quad W_i(\mathbf{e}_i, t_i + \Delta_{i-1}) + R_i(m). \tag{4.36}$$

By subtracting (4.35) from (4.36), we get

$$q_n - q_1 = D^i_{\Delta_{i-1}}(\mathbf{e}_i, t_i) + R_i(m). \tag{4.37}$$

Further notice that

$$\Delta_i - \Delta_{i-1} = (n-1) \times (G_i(m) - G_{i-1}(m)) = q_n - q_1. \tag{4.38}$$

Combining (4.37) and (4.38), the second part of (4.31) follows. ∎

We assume that adjacent packet-trains in the probing series are sufficiently separated so that the transient probing intrusion effect on each queue caused by the previous packet-train is assimilated before the arrival of the next packet-train. Consequently, each packet-train comes to see a multi-hop system of the same stochastic nature and the output dispersion process $\{G_N(m), 1 \le m < \infty\}$ is a *identically distributed* random process.

By further assuming the existence of a sample-path time-average, we have that with probability one, the asymptotic average of the packet train output

---

[3]For the details about "intrusion residual", please refer to [23]

dispersions coincide with the ensemble mean of (any random variable in) the stationary process $\{G_N(m), 1 \leq m < \infty\}$. Therefore, in the rest of the chapter, we focus on the statistics of the output dispersion random variable $G_N$, where we drop the index $m$.

## 4.4  Multi-Hop Response Curves

In this section, we first apply mathematical induction to show that the gap response curve $\mathcal{Z} = E[G_N(g_I, s, n)]$ of a multi-hop path $\mathcal{P}$ is lower bounded by its CRF counterpart $\mathcal{F} = \gamma_N(g_I, s)$. We then investigate the impact of packet-train parameters on probing response.

### 4.4.1  Bound

Our next lemma shows that passing through a link can only increase the ensemble mean of the dispersion random variable.

**Lemma 12** *For $1 \leq i \leq N$, the output dispersion random variable $G_i$ has a mean no less than that of $G_{i-1}$. That is, $E[G_i] \geq E[G_{i-1}]$.*

**Proof:**  First, due to the second part of (4.31), we have

$$E[G_i] = E[G_{i-1}] + \frac{E[D^i_{\Delta_{i-1}}(\mathbf{e}_i, t_i)]}{n-1} + \frac{E[R_i]}{n-1}. \tag{4.39}$$

Note that the second term in the right hand side of (4.39) is 0, regardless of the distribution of the random variable $\Delta_{i-1}$. Let $P(x)$ be the distribution function of $\Delta_{i-1}$, we have

$$E[D^i_{\Delta_{i-1}}(\mathbf{e}_i, t_i)] = \int_0^\infty E[D^i_{\Delta_{i-1}}(\mathbf{e}_i, t_i)|\Delta_{i-1} = x]dP(x)$$
$$= \int_0^\infty E[D^i_x(\mathbf{e}_i, t_i)]dP(x) = \int_0^\infty 0 \quad dP(x) = 0. \tag{4.40}$$

Also note that due to the properties of the intrusion residual, $E[R_i] \geq 0$. Hence $E[G_i] \geq E[G_{i-1}]$. ∎

Using the first part of (4.31), our next lemma shows that for any link $L_i$, the output dispersion random variable $G_i$ is lower bounded in mean by a linear combination of the output dispersion random variables $G_k$, where $k < i$.

**Lemma 13** *For $1 \leq i \leq N$, the output dispersion random variable $G_i$ satisfies the following inequality*

$$E[G_i] \geq \frac{1}{C_i} \left( \sum_{k=1}^{i} \Psi^T \Gamma_i^k E[G_{k-1}] + s \right). \tag{4.41}$$

**Proof:** Due to the first part of (4.31), we have

$$E[G_i] = \frac{1}{C_i} \left( \sum_{k=1}^{i} E\left[ Y_{\Delta_{k-1}}^k (\Gamma_i^k, t_k) G_{k-1} \right] + s \right) + \frac{E[\tilde{I}_i]}{n-1}. \tag{4.42}$$

Let $P(x)$ be the distribution function of $\Delta_{k-1}$, we have

$$
\begin{aligned}
& E\left[ Y_{\Delta_{k-1}}^k (\Gamma_i^k, t_k) \times G_{k-1} \right] \\
&= \int_0^\infty E\left[ Y_{\Delta_{k-1}}^k (\Gamma_i^k, t_k) \times G_{k-1} \middle| G_{k-1} = \frac{x}{n-1} \right] dP(x) \\
&= \int_0^\infty E\left[ Y_x^k (\Gamma_i^k, t_k) \times \frac{x}{n-1} \right] dP(x) \\
&= \frac{\Psi^T \Gamma_i^k}{n-1} \int_0^\infty x \, dP(x) = \Psi^T \Gamma_i^k E[G_{k-1}]. \tag{4.43}
\end{aligned}
$$

Combining (4.42), (4.43), and the fact that $E[\tilde{I}_i] \geq 0$, the lemma follows. ∎

Combining Lemma 12 and Lemma 13, we get

$$E[G_i] \geq \max \left( E[G_{i-1}], \frac{\sum_{k=1}^{i} \Psi^T \Gamma_i^k E[G_{k-1}] + s}{C_i} \right). \tag{4.44}$$

This leads to the following theorem.

**Theorem 11** *For any input dispersion $g_I$, packet-train parameters $s$ and $n$, the output dispersion random variable $G_N$ of path $\mathcal{P}$ is lower bounded in mean by the output dispersion $\gamma_N(g_I, s)$ on the CRF counterpart of $\mathcal{P}$. That is*

$$E[G_N(g_I, s, n)] \geq \gamma_N(g_I, s). \tag{4.45}$$

**Proof:** We apply mathematical induction to $i$. When $i = 0$, $E[G_0(g_I, s, n)] = \gamma_0(g_I, s) = g_I$. Assuming that (4.45) holds for $0 \leq i < N$, we next prove that it also holds for $i = N$. Recall (4.44), we have

$$E[G_N] \geq \max\left( E[G_{N-1}], \frac{\sum_{k=1}^{N} \Psi^T \Gamma_N^k E[G_{k-1}] + s}{C_N} \right)$$

$$\geq \max\left( \gamma_{N-1}, \frac{\sum_{k=1}^{N} \Psi^T \Gamma_N^k \gamma_{k-1} + s}{C_N} \right) = \gamma_N, \tag{4.46}$$

where the second inequality is due to induction hypothesis, and the last equality is due to Theorem 10. ∎

Theorem 11 shows that in the entire input gap range, the piece-wise linear CRF gap response curve $\mathcal{F}$ discussed in Section 4.2 is a lower bound of the real gap curve $\mathcal{Z}$. The deviation between the real curve $\mathcal{Z}$ and its fluid lower bound $\mathcal{F}$, which is denoted by $\beta_N(g_I, s, n)$, can be recursively expressed as follows

$$\beta_i = \begin{cases} \beta_{i-1} + \dfrac{E[R_i]}{n-1} & L_i \notin \mathcal{H}(\overline{\mathcal{P}}, s/g_I) \\ \dfrac{1}{C_i} \sum_{k=1}^{i} \Psi^T \Gamma_i^k \beta_{k-1} + \dfrac{E[\tilde{I}_i]}{n-1} & L_i \in \mathcal{H}(\overline{\mathcal{P}}, s/g_I) \end{cases}, \tag{4.47}$$

where $\overline{\mathcal{P}}$ is the CRF counterpart of path $\mathcal{P}$, and $\beta_0 = 0$. Expanding (4.47), we get the following result.

**Corollary 6** *The response curve deviation $\beta_N(g_I, s, n)$ has the following non-recursive expression*

$$\beta_N = \frac{1}{n-1} \left( \sum_{L_j \notin \mathcal{H}} a_j E[R_j] + \sum_{L_k \in \mathcal{H}} b_k E[\tilde{I}_k] \right), \tag{4.48}$$

where $\mathcal{H} = \mathcal{H}(\overline{\mathcal{P}}, s/g_I)$, $a_j$ and $b_k$ are coefficients that are only functionally related to the routing matrix $\mathbf{R}$, the flow rate vector $\Psi$, and the capacity vector $\mathbf{C}$, but are not functionally related to the input packet-train parameters $s$ and $n$.

In what follows, we study the asymptotics of the curve deviation $\beta_N$ when input packet-train parameters $s$ and $n$ become large and show that the fluid lower bound $\mathcal{F}$ is in fact a *tight* bound of the real response curve $\mathcal{Z}$.

### 4.4.2  Impact of Probing Packet Size

We now demonstrate that for any input probing rate $r_I$, the curve deviation $\beta(s/r_I, s, n)$ vanishes as probing packet size $s$ approaches infinite. We prove this result under the condition of one-hop persistent cross-traffic routing. We also justify this conclusion informally for arbitrary cross-traffic routing and point out the major difficulty for a rigorous proof. First, we make an additional assumption as follows.

**Assumption 5** *Denote by $P_\delta^i(x)$ the distribution function of the $\delta$-interval available bandwidth process $\{B_\delta^i(\mathbf{e}_i, t)\}$, we assume for all $1 \leq i \leq N$, the following holds*

$$\begin{cases} P_\delta^i(r) = o\left(\dfrac{1}{\delta^2}\right) & r < C_i - \Psi^T \mathbf{e}_i \\ P_\delta^i(r) = 1 - o\left(\dfrac{1}{\delta^2}\right) & r > C_i - \Psi^T \mathbf{e}_i \end{cases}. \tag{4.49}$$

Recall that the mean-square ergodicity assumption we made earlier implies that as the observation interval $\delta$ gets large, the random variable $B_\delta^i(\mathbf{e}_i, t)$ converges in distribution to $C_i - \Psi^T \mathbf{e}_i$. Assumption 5 further ensures that this convergence is *fast* in the sense of (4.49). Even though this condition appears

cryptic at first, it is valid in a broad range of cross-traffic environments. Next, we show its validity under the condition of regenerative [4] link utilization.

**Theorem 12** *When hop utilization process $\{U_i(\mathbf{e}_i, t)\}$ is regenerative, condition (4.49) holds.*

**Proof:** When the hop utilization process $\{U_i(\mathbf{e}_i, t)\}$ is regenerative, the process $\{C_i(1 - U_i(\mathbf{e}_i, t))\}$ is also regenerative with the same stopping times and regeneration cycles. Further note that the $\delta$-interval available bandwidth $B_\delta^i(\mathbf{e}_i, t)$ is the time average of the regenerative process $\{C_i(1 - U_i(\mathbf{e}_i, t))\}$. According to the regenerative central limit theorem [42, pages 124], for any $t$, $B_\delta^i(\mathbf{e}_i, t)$ converges in distribution to a Gaussian random variable $N(C_i - \Psi^T \mathbf{e}_i, \sigma^2/\delta)$ as $\delta$ approaches infinity, where $\sigma$ is a constant. This implies that the mean of the Gaussian distribution remains $C_i - \Psi^T \mathbf{e}_i$ for all $\delta$ while the variance is inversely proportional to $\delta$. Therefore, for sufficiently large $\delta$, we have

$$P_\delta^i(r) = \frac{1}{2}\left(1 + \mathbf{erf}\left(\frac{(r - C + \Psi^T \mathbf{e}_i)\sqrt{\delta}}{\sigma\sqrt{2}}\right)\right), \tag{4.50}$$

where **erf** is the special function called *Gauss error function*.

According to the asymptotic series of $\mathbf{erf}(x)$ [3, pages 297-309], we have

$$\mathbf{erf}(x) = \begin{cases} \Theta\left(\frac{-1}{xe^{x^2}}\right) - 1 & x < 0 \\ \Theta\left(\frac{-1}{xe^{x^2}}\right) + 1 & x > 0 \end{cases}. \tag{4.51}$$

Combining (3.100) with (4.50), we have

$$P_\delta^i(r) = \begin{cases} \Theta\left(\frac{1}{\sqrt{\delta}e^{k\delta}}\right) = o\left(\frac{1}{\delta^2}\right) & r < C_i - \Psi^T \mathbf{e}_i \\ 1 - \Theta\left(\frac{1}{\sqrt{\delta}e^{k\delta}}\right) = 1 - o\left(\frac{1}{\delta^2}\right) & r > C_i - \Psi^T \mathbf{e}_i \end{cases}, \tag{4.52}$$

---

[4]Refer to [42, pages 89] for the definition of regenerative process.

where $k$ is a positive constant given below

$$k = \frac{\left(r - \Psi^T \mathbf{e}_i\right)^2}{2\sigma^2}. \tag{4.53}$$

This proves the theorem, at the same time, (4.52) shows that the convergence is much faster than what we assumed in Assumption 5. It is in fact exponentially fast. ∎

Regenerative queue is very common both in practice and in stochastic modeling literature. Note that all the four traffic types used in [23] lead to regenerative hop workload and consequently lead to regenerative link utilization. We also conjecture that (4.49) holds under much milder condition; and we leave its identification as future work.

An immediate consequence of Assumption 5 is the following lemma.

**Lemma 14** *For any link $L_i$ in $\mathcal{P}$, assuming $\mathbf{e}_i = \mathbf{r}_i$, when $L_i$ is probed by packet-pairs with input rate $r$, we have the follow two limits regarding the conditional second-order moments of $R_i$ and $\tilde{I}_i$.*

$$\begin{cases} \lim_{s \to \infty} E[R_i^2 | G_{i-1} = s/r] = 0 & r < C_i - \lambda_i \\ \lim_{s \to \infty} E[\tilde{I}_i^2 | G_{i-1} = s/r] = 0 & r > C_i - \lambda_i \end{cases}. \tag{4.54}$$

**Proof:** We first consider the case when $r < C_i - \lambda$. Let $\delta = s/r$ and denote by $R_i(\delta)$ the random variable $R_i$ under the condition that the input packet-pair dispersion $G_{i-1} = \delta$. We have

$$R_i(\delta) = R_i(s/r) = \max\left(0, \frac{s - \delta B_\delta^i(\mathbf{e}_i, t)}{C_i}\right), \tag{4.55}$$

where $t$ is the arrival time of the packet-pair into $L_i$. Denoting by $P_\delta^i(x)$ the distribution function of the random variable $B_\delta^i(\mathbf{e}_i, t)$, we have

$$E[R_i^2(\delta)] = \int_0^r \frac{\delta^2(r-x)^2}{C_i^2} dP_\delta^i(x) \leq \frac{2r^2\delta^2 P_\delta^i(r)}{C_i^2}. \tag{4.56}$$

Taking the limit of (4.56) and further recalling Assumption 5, we get

$$0 \leq \lim_{\delta \to \infty} E[R_i^2(\delta)] \leq \lim_{\delta \to \infty} \frac{2r^2\delta^2 P_\delta(r)}{C_i^2} = 0. \tag{4.57}$$

This leads to the first part in (4.54). Now consider the case when $r > C_i - \lambda_i$. Denoting by $\tilde{I}_i(\delta)$ the random variable $\tilde{I}_i$ under the condition that the input packet-pair dispersion $G_{i-1} = \delta = s/r$, we have

$$\tilde{I}_i(\delta) = \tilde{I}_i(s/r) = \max\left(0, \frac{\delta B_\delta^i(\mathbf{e}_i, t) - s}{C_i}\right). \tag{4.58}$$

Computing the second moment of $\tilde{I}_i(s/r)$, we get

$$E[\tilde{I}_i^2(\delta)] = \int_r^C \frac{\delta^2(x-r)^2}{C_i^2} dP_\delta^i(x) \leq \frac{\delta^2(C_i-r)^2}{C_i^2}\left(1 - P_\delta^i(r)\right). \tag{4.59}$$

Taking the limit of (4.59) and recalling Assumption 5, we get

$$0 \leq \lim_{\delta \to \infty} E[\tilde{I}_i^2(\delta)] \leq \lim_{\delta \to \infty} \frac{(C_i-r)^2\delta^2(1 - P_\delta^i(r))}{C_i^2} = 0. \tag{4.60}$$

This leads to the second part in (4.54). ∎

Our next theorem states formally the convergence property of the output dispersion random variable $G_N(s/r_I, s, n)$ when $s$ increases, for any $N$-hop path $\mathcal{P}$ with one-hop persistence cross-traffic routing, for any given input rate $r_I \in (0, \infty)$, and any packet-train length $n \geq 2$.

**Theorem 13** *Given one-hop persistent cross-traffic routing and the three assumptions made in this chapter, for any input rate $r_I$, the output dispersion random variable $G_N$ of path $\mathcal{P}$ converges in mean to its fluid lower bound $\gamma_N$:*

$$\lim_{s \to \infty} E[G_N(s/r_I, s, n)] = \gamma_N(s/r_I, s). \tag{4.61}$$

*The asymptotic variance of $G_N$ when $s$ increases is upper bounded by some constant $K_N$,*

$$\lim_{s \to \infty} E[(G_N - \gamma_N)^2] \leq K_N. \tag{4.62}$$

**Proof:** We only consider the case of packet-pair probing. The proof can be easily extended to packet-train probing by applying mathematical induction on $n$. In the proof of packet-pair case, we apply mathematical induction on $i$.

For the base case when $i = 0$, $G_0 = s/r_I = \gamma_0$ and $K_0 = 0$, the theorem holds trivially. Suppose the theorem holds for $i = N - 1$, we now show that it also holds for $i = N$.

First consider the case when $s/\gamma_{N-1} < C_N - \lambda_N$, due to Lemma 11, we have:

$$G_N = G_{N-1} + D^N_{G_{N-1}}(\mathbf{e}_N, t) + R_N. \tag{4.63}$$

We now examine the asymptotic mean and asymptotic variance for each of the three terms on the right hand side of (4.63). For the first term $G_{N-1}$, due to the induction hypothesis, we have

$$\lim_{s \to \infty} E[G_{N-1}] = \gamma_{N-1}, \tag{4.64}$$

$$\lim_{s \to \infty} E[(G_{N-1} - \gamma_{N-1})^2] \le K_{N-1}. \tag{4.65}$$

The second term in (4.63) is a zero-mean random variable regardless of the distribution of $G_{N-1}$,

$$\lim_{s \to \infty} E\left[D^N_{G_{N-1}}(\mathbf{e}_N, t)\right] = 0. \tag{4.66}$$

The proof is similar to what we showed in (4.40). The variance of $D^N_{G_{N-1}}(\mathbf{e}_N, t)$ converges to $2Var[W_N(\mathbf{e}_N, t)]$ as $s \to \infty$, which is a constant with respect to $s$. To show this, first note that

$$E\left[\left(D^N_{G_{N-1}}(\mathbf{e}_N, t)\right)^2\right] = \int_0^\infty \left(D^N_x(\mathbf{e}_N, t)\right)^2 dP(x), \tag{4.67}$$

where $P(x)$ is the distribution function of $G_{N-1}$. The integral term in (4.67) can be decomposed into the sum of three integral terms as follows:

$$E[D^2] = \int_0^{E/2} D^2 dP(x) + \int_{E/2}^{3E/2} D^2 dP(x) + \int_{3E/2}^\infty D^2 dP(x), \tag{4.68}$$

where $D^2 = \left(D_x^N(\mathbf{e}_N, t)\right)^2$ and $E = E[G_{N-1}]$. Using Chebyshev's inequality and the fact that $E[D^2] \leq 4Var[W_N(\mathbf{e}_N, t)]$, it is easy to show that both the first and the third integral terms in (4.68) converges to 0 as $s \to \infty$. In addition, using Chebyshev's inequality and lemma 10, we can show that the second integral term in (4.68) converges to $2Var[W_N(\mathbf{e}_N, t)]$ as $s \to \infty$. Omitting all the intermediate steps, we get:

$$\lim_{s \to \infty} E\left[\left(D_{G_{N-1}}^N(\mathbf{e}_N, t)\right)^2\right] = 2Var[W_N(\mathbf{e}_N, t)]. \tag{4.69}$$

For the third term $R_N$ in (4.63), its first-order moment converges to 0 as $s \to \infty$ as we show next. Note that

$$E[R_N] = \int_0^{s/A_N} E[R_N(x)]dP(x) + \int_{s/A_N}^{\infty} E[R_N(x)]dP(x), \tag{4.70}$$

where $A_N = C_N - \lambda_N$ is the available bandwidth of $L_N$, $P(x)$ is the distribution function of $G_{N-1}$, and $E[R_N(x)]$ denotes the conditional expectation $E[R_N|G_{N-1} = x]$. Notice that $R_N$ is upper bounded by $s/C_N$, Hence due to Chebyshev's inequality, for the first additive term in (4.70), we have

$$
\begin{aligned}
0 \leq \int_0^{s/A_N} E[R_N(x)]dP(x) &\leq \frac{s}{C_N}P\left(\frac{s}{A_N}\right) \\
&\leq \frac{sVar[G_{N-1}]}{C_N\left(s/A_N - E[G_{N-1}]\right)^2}.
\end{aligned}
\tag{4.71}
$$

Taking the limit of (4.71) when $s \to \infty$, we get

$$
\begin{aligned}
0 &\leq \lim_{s \to \infty} \int_0^{s/A_N} E[R_N(x)]dP(x) \\
&\leq \lim_{s \to \infty} \frac{s}{C_N} \frac{Var[G_{N-1}]}{\left(s/A_N - E[G_{N-1}]\right)^2} \\
&\leq \lim_{s \to \infty} \frac{s}{C_N} \frac{K_{N-1}}{\left(s/A_N - \gamma_{N-1}\right)^2} = \lim_{s \to \infty} \Theta\left(\frac{1}{s}\right) = 0,
\end{aligned}
\tag{4.72}
$$

where the last inequality is due to induction hypothesis and the second last equality is due to the fact that $\gamma_{N-1}$ is a linear function of $s/r_I$ as stated in Property 1.

For the second additive term in (4.70), first recall Theorem 6, which says that $R_N(x)$ is a monotone decreasing function of $x$. Therefore, we have

$$
\begin{aligned}
0 &\leq \lim_{s \to \infty} \int_{s/A_N}^{\infty} E[R_N(x)] dP(x) \\
&\leq \lim_{s \to \infty} E[R_N(s/A_N)] = 0,
\end{aligned}
\tag{4.73}
$$

where the last equality is due to Lemma 14. From Lemma 14, it follows that $R_N(s/A_N)$ converges in mean-square sense to 0, which implies that $R_N(s/A_N)$ also converges to 0 in mean when $s \to \infty$. Combing (4.72) and (4.73), it follows that

$$
\lim_{s \to \infty} E[R_N] = 0.
\tag{4.74}
$$

In almost the same way as showed from (4.70) to (4.74), we can prove that the asymptotic variance of $R_N$ when $s$ increases is bounded by a constant. We omit the proof details of this step. Combining all these investigation, it follows that

$$
\lim_{s \to \infty} E[G_N] = \lim_{s \to \infty} E[G_{N-1}] = \gamma_{N-1} = \gamma_N.
\tag{4.75}
$$

The asymptotic variance of $G_N$ is also bounded by a constant irrespective of $s$ due to the fact that all the three additive terms on the right hand side of (4.63) have so bounded asymptotic variance. We denote this variance upper bound by $K_N$.

So far, we finished the proof for the case when $s/\gamma_{N-1} < A_N$. For the case when $s/\gamma_{N-1} > A_N$, we have the following due to Lemma 11 and the one-hop persistent routing assumption:

$$
G_N = \frac{Y_{G_{N-1}}^N(\mathbf{e}_N, t) G_{N-1}}{C_N} + \frac{s}{C_N} + \tilde{I}_N
\tag{4.76}
$$

We now examine the asymptotic mean and variance for each of the additive terms on the right hand side of (4.76). For the first term we have

$$
\begin{aligned}
\lim_{s \to \infty} E &\left[ \frac{Y^N_{G_{N-1}}(\mathbf{e}_N, t) G_{N-1}}{C_N} \right] \\
&= \frac{1}{C_N} \lim_{s \to \infty} \int_0^\infty E \left[ Y^N_{G_{N-1}}(\mathbf{e}_N, t) G_{N-1} \Big| G_{N-1} = x \right] dP(x) \\
&= \frac{\Psi^T \mathbf{e}_N}{C_N} \lim_{s \to \infty} \int_0^\infty x \, dP(x) = \frac{\lambda_N \gamma_N}{C_N},
\end{aligned}
\tag{4.77}
$$

where $P(x)$ is the distribution function of $G_{N-1}$. Similarly, we can get the asymptotic variance as follows

$$
\begin{aligned}
\lim_{s \to \infty} Var &\left[ \frac{Y^N_{G_{N-1}}(\mathbf{e}_N, t) G_{N-1}}{C_N} \right] \\
&= \frac{\lambda_N^2}{C_N^2} \lim_{s \to \infty} Var[G_{N-1}] \le \frac{\lambda_N^2 K_{N-1}}{C_N^2},
\end{aligned}
\tag{4.78}
$$

where the last inequality is due to induction hypothesis. Note that the limiting variance is bounded by a constant irrelevant to $s$.

The second additive term in (4.76) is a constant. For the third term $\tilde{I}_N$, we now show that it converges to 0 in mean-square sense as $s \to \infty$. Consequently, both the asymptotic mean and the asymptotic variance of this term is 0. Note that $E[\tilde{I}_N^2]$ can be decomposed as

$$
E\left[ \tilde{I}_N^2 \right] = \int_0^{s/A_N} E\left[ \tilde{I}_N^2(x) \right] dP_s(x) + \int_{s/A_N}^\infty E\left[ \tilde{I}_N^2(x) \right] dP_s(x),
\tag{4.79}
$$

where $A_N = C_N - \lambda_N$ is the available bandwidth of $L_N$, $P_s(x)$ is the distribution function of $G_{N-1}$ given packet size $s$, and $E[\tilde{I}_N^2(x)]$ denotes the conditional second moment $E[\tilde{I}_N^2 | G_{N-1} = x]$. Note that the first term in (4.79) approaches

0 as $s \to \infty$. That is,

$$
\begin{aligned}
\lim_{s \to \infty} \int_0^{s/A_N} E\left[\tilde{I}_N^2(x)\right] dP_s(x) \\
= \lim_{s \to \infty} \int_{A_N}^{\infty} E\left[\tilde{I}_N^2\left(\frac{s}{r}\right)\right] d\tilde{P}_s(r) \\
\leq \lim_{s \to \infty} E\left[I_N^2\left(\frac{s}{A_N}\right)\right] = 0,
\end{aligned}
\tag{4.80}
$$

where $\tilde{P}_s(r)$ is the distribution function of the random variable $s/G_{N-1}$ given $s$ fixed. The inequality is due to fact that $\tilde{I}_N(x)$ is a monotone decreasing function of $x$ given $s$ fixed, as stated in Theorem 6. The last equality in (4.80) is due to Lemma 14.

The second term in (4.79) also approaches 0 as $s \to \infty$. Note that $\tilde{I}_N(x) \leq x$, so we have

$$
\begin{aligned}
\lim_{s \to \infty} \int_{s/A_N}^{\infty} E\left[\tilde{I}_N^2(x)\right] dP_s(x) \leq \lim_{s \to \infty} \int_{s/A_N}^{\infty} x^2 dP_s(x) \\
= \lim_{s \to \infty} \left( E[G_{N-1}^2(s)] - \int_0^{s/A_N} x^2 dP_s(x) \right) \\
= \lim_{s \to \infty} E[G_{N-1}^2(s)] - \lim_{s \to \infty} E[G_{N-1}^2(s)] = 0.
\end{aligned}
\tag{4.81}
$$

Combining (4.80) and (4.81), we get

$$
\lim_{s \to \infty} E[\tilde{I}_N^2] = \lim_{s \to \infty} E[\tilde{I}_N] = 0.
\tag{4.82}
$$

Combining (4.82) and (4.77), we get

$$
\lim_{s \to \infty} E[G_N] = \frac{\lambda_N \gamma_{N-1} + s}{C_N} = \gamma_N.
\tag{4.83}
$$

Combining induction hypothesis, (4.78), and (4.83), we get

$$
\lim_{s \to \infty} Var[G_N] \leq \frac{\lambda_N^2}{C_N^2} \lim_{s \to \infty} Var[G_{N-1}] = K_N,
\tag{4.84}
$$

which is a constant irrelevant to $s$. Combining the two cases, we complete the inductive step for any probing input rate $r_I$. Hence, the theorem follows. ∎

Note that the bounded variance, as stated in (4.62), is not an extra result in addition to the mean convergence given by (4.61). Rather, it is an inseparable part of the whole theorem. Without bounded variance of the output dispersion random variable $G_{N-1}$, we can not obtain the mean convergence of $G_N$ to $\gamma_N$.

We further point that by assuming one-hop persistent cross-traffic routing, we have avoided analyzing the departure processes of cross-traffic flows. When a traversing flow $f_j$ of link $L_i$ enters the path from some upstream link of $L_i$, the arrival process of $f_j$ at $L_i$ is its departure process at $L_{i-1}$, which is usually not the same as that at the entrance link of $f_j$. Unfortunately, in the queueing theory literature, there is no exact result for departure processes in FCFS queueing network models if one goes beyond the assumption of Poisson-arrival. Motivated by the intractability of the models, researchers have focused their attentions on approximations [30][24].

To prove Theorem 13 in arbitrary cross-traffic routing, we also need an approximation assumption which says that any cross-traffic flow (and consequently flow aggregation) that traverses link $L_i$ (regardless wether it enters the path from $L_i$ or some upstream link of $L_i$) exhibits ergodic stationary arrival at $L_i$. Under this assumption, which we call "stationary departure approximation", it becomes easy to show that Theorem 13 holds for arbitrary cross-traffic routing. We skip the details of this extension and next apply the stationary departure approximation to examine the impact of packet-train length $n$.

### 4.4.3  Impact of Packet-Train Length

We now show that when packet-size $s$ is kept constant, as the packet-train length $n \to \infty$, the output dispersion random variable $G_N(g_I, s, n)$ of path $\mathcal{P}$ converges

in mean-square sense to its fluid lower bound $\gamma_N(g_I, s)$, for any $g_I$ and any $s$. This means that not only $E[G_N]$ converges to $\gamma_N$, but also the variance of $G_N$ decays to 0 as $n$ increases. We first prove this result over a single-hop path. We then apply mathematical induction to extend this conclusion to any multi-hop path with arbitrary cross-traffic routing under the assumption of "stationary departure approximation".

**Theorem 14** *Under the first assumption of this chapter, for a single-hop path $\mathcal{P}$ with capacity $C$ and cross-traffic intensity $\lambda < C$, for any input dispersion $g_I \in (0, \infty)$ and probing packet size $s$, the output dispersion random variable $G$ converges to its fluid lower bound $\gamma$ in mean-square sense as $n \to \infty$,*

$$\lim_{n \to \infty} E\left[\left(G(g_I, s, n) - \max\left(g_I, \frac{\lambda g_I + s}{C}\right)\right)^2\right] = 0. \qquad (4.85)$$

**Proof:** First consider the case when $s/g_I < C - \lambda$. We first examine the output sampling interval random variable $\Delta = (n-1)G$. The key is to view the first and last packets in the input packet-train as a packet-pair and view the other packets in between as if they were from another cross-traffic flow $f'$. The real cross-traffic and $f'$ together form a flow aggregation denoted by $\mathbf{p}$. Obviously, the packet arrival in $\mathbf{p}$ is still ergodic stationary. The long term arrival rate of $\mathbf{p}$ is $\lambda + s/g_I < C$. The workload-difference process $D_\delta(\mathbf{p}, t)$ is a zero-mean process. According to Lemma 11, $\Delta$ can be expressed as follows

$$\Delta = (n-1)g_I + D_\delta(\mathbf{p}, t) + R, \qquad (4.86)$$

where $t$ is the arrival time of the first probing packet into the hop, $\delta = (n-1)g_I$ is the sampling interval of the input packet-train, $R = \max(0, s - B_\delta(\mathbf{p}, t)\delta)$ is the intrusion residual with respect to the flow aggregation $\mathbf{p}$. The output

dispersion $G = \Delta/(n-1)$ can be expressed as

$$G = g_I + \frac{D_\delta(\mathbf{p},t)}{n-1} + \frac{\max(0, s - B_\delta(\mathbf{p},t)\delta)}{n-1}, \qquad (4.87)$$

Notice that, as $n$ increases, the second additive term converges to 0 in mean-square sense. That is,

$$\lim_{n\to\infty} E\left[\left(\frac{D_\delta(\mathbf{p},t)}{n-1}\right)^2\right] = \frac{2Var[W(\mathbf{p},t)]}{\lim_{n\to\infty}(n-1)^2} = 0, \qquad (4.88)$$

where the first equality is due to Lemma 10. The third term on the right hand side of (4.87) also converge to 0 in mean-square sense.

$$\lim_{n\to\infty} E\left[\left(\frac{\max(0, s - B_\delta(\mathbf{p},t)\delta)}{n-1}\right)^2\right] \le \lim_{n\to\infty} \frac{s^2}{(n-1)^2} = 0. \qquad (4.89)$$

Combining (4.87), (4.88), and (4.89), we get

$$\lim_{n\to\infty} E\left[(G(g_I, s, n) - g_I)^2\right] = 0. \qquad (4.90)$$

Now consider the case when $s/g_I > C - \lambda$. We still examine the sampling interval interval $\Delta$, and according to Lemma 11, we have

$$\Delta = \frac{Y_\delta(\mathbf{p},t)\delta}{C} + \frac{s}{C} + \tilde{I}, \qquad (4.91)$$

where $t$ and $\delta$ have the same meanings as those in (4.87). $\tilde{I} = \max\left(0, B_\delta(\mathbf{p},t)\delta - s\right)$ is the hop idle time during the interval $[t, t + \Delta]$. The output dispersion $G = \Delta/(n-1)$ can be expressed as

$$G = \frac{Y_\delta(\mathbf{p},t)\delta}{(n-1)C} + \frac{s}{(n-1)C} + \frac{\max\left(0, B_\delta(\mathbf{p},t)\delta - s\right)}{n-1}. \qquad (4.92)$$

The first additive term in (4.92) converges in mean-square sense to $(\lambda g_I + s)/C$, showed in the following:

$$\lim_{n\to\infty} E\left[\left(\frac{Y_\delta(\mathbf{p},t)\delta - (n-1)(\lambda g_I + s)}{(n-1)C}\right)^2\right]$$

$$= \frac{g_I^2}{C^2} \lim_{\delta\to\infty} E\left[\left(Y_\delta(\mathbf{p},t) - \left(\lambda + \frac{s}{g_I}\right)\right)^2\right] = 0, \qquad (4.93)$$

where the second equality is due to the mean-square ergodicity of the flow aggregation $\mathbf{p}$. The second term in (4.92) is deterministic, and its square converges to 0 as $n \to \infty$. The third term in (4.92) converges in mean-square sense to 0 when $n$ increases. To show that, first notice that since the arrival rate of $\mathbf{p}$ is greater than hop capacity $C$, we have

$$\lim_{\delta \to \infty} E[B_\delta(\mathbf{p}, t)] = 0. \tag{4.94}$$

further notice that $B_\delta(\mathbf{p}, t)$ is distributed in a *finite* interval $[0, C]$. Hence, (4.93) implies that the second moment of $B_\delta(\mathbf{p}, t)$ also converges to 0 as $\delta$ increases,

$$\lim_{\delta \to \infty} E\left[(B_\delta(\mathbf{p}, t))^2\right] = 0. \tag{4.95}$$

This leads to the following

$$
\begin{aligned}
0 \leq \lim_{n \to \infty} E&\left[\left(\frac{\max\left(0, B_\delta(\mathbf{p}, t)\delta - s\right)}{n - 1}\right)^2\right] \\
&\leq \lim_{n \to \infty} E\left[\left(\frac{B_\delta(\mathbf{p}, t)\delta}{n - 1}\right)^2\right] \\
&= \lim_{\delta \to \infty} g_I^2 E\left[(B_\delta(\mathbf{p}, t))^2\right] = 0.
\end{aligned}
\tag{4.96}
$$

Combining (4.92), (4.93), and (4.96), we get

$$\lim_{n \to \infty} E\left[\left(G(g_I, s, n) - \frac{\lambda g_I + s}{C}\right)^2\right] = 0. \tag{4.97}$$

Combining (4.90) and (4.97), the theorem follows. ∎

Our next theorem extends this result to multi-hop path with arbitrary cross-traffic routing.

**Theorem 15** *Under the first two assumptions of this chapter[5] and the " stationary departure approximation", for any $N$-hop path $\mathcal{P}$ with arbitrary cross-traffic routing, for any input dispersion $g_I \in (0, \infty)$ and probing packet size $s$,*

---

[5]Note that Assumption 5 is not necessary in this theorem.

*the output dispersion random variable $G_N$ converges to its fluid lower bound $\gamma_N$ in mean-square sense as $n \to \infty$,*

$$\lim_{n \to \infty} E\left[(G_N(g_I, s, n) - \gamma_N(g_I, s))^2\right] = 0. \tag{4.98}$$

**Proof:** We apply induction on $i$. When $i = 1$, the conclusion holds due to Theorem 14. Assuming that (4.98) holds for all $i < N$, we next show it also holds for $i = N$.

We apply the same trick as we do in the proof of Theorem 14. We view the first and last probing packets $p_1$ and $p_n$ as a packet-pair, and view the rest of probing packets in the train as if they were from another cross-traffic flow $f'$. We denote the aggregation of $\mathbf{r}_N$ and $f'$ as $\mathbf{p}$. Due to the "stationary departure approximation", the traffic arrival in $\mathbf{p}$ can be viewed as ergodic stationary when $n$ is sufficient large. We now examine the average arrival rate of $\mathbf{p}$ at link $L_N$. That is, we compute

$$\lambda_{\mathbf{p}} = \lim_{n \to \infty} \frac{E[\Omega_N]}{(n-1)E[G_{N-1}(g_I, s, n)]} \tag{4.99}$$

where $\Omega_N$ is the random variable indicating the volume of traffic buffered between $p_1$ and $p_n$ in the outgoing queue of $L_N$. Notice that

$$E[\Omega_N] = E\left[\sum_{k=1}^{N} Y_{\Delta_{k-1}}(\Gamma_N^k, t_k)\Delta_{k-1}\right] + (n-1)s, \tag{4.100}$$

where $t_k$ is the arrival time of $p_1$ at $L_k$, $\Delta_{k-1} = (n-1)G_{k-1}$ is the sampling interval of the input packet-pair $p_1$ and $p_n$ at $L_k$. Substitute (4.100) back into (4.99), we get the following due to induction hypothesis:

$$\begin{aligned}
\lambda_{\mathbf{p}} &= \lim_{n \to \infty} \frac{\sum_{k=1}^{N} E[Y_{\Delta_{k-1}}(\Gamma_N^k, t_k)G_{k-1}] + s}{E[G_{N-1}(g_I, s, n)]} \\
&= \frac{\sum_{k=1}^{N} \Psi^T \Gamma_N^k \gamma_{k-1} + s}{\gamma_{N-1}}.
\end{aligned} \tag{4.101}$$

We now consider the case when $\lambda_{\mathbf{p}} < C_N$. This leads to $\gamma_N = \gamma_{N-1}$ due to Theorem 10 and (4.101). Further, due to Lemma 11, we have

$$\Delta_N = \Delta_{N-1} + D^N_{\Delta_{N-1}}(\mathbf{p}, t) + R_N, \tag{4.102}$$

where $t$ is the arrival time of $p_1$ at $L_N$, and $R_N = \max(0, s - B^N_{\Delta_{N-1}}(\mathbf{p}, t)\Delta_{N-1})$ is the intrusion residual of $p_1$ on $p_n$ with respect to $W_N(\mathbf{p}, t)$. Dividing $n-1$ at both sides of (4.102), we get

$$G_N = G_{N-1} + \frac{D^N_{\Delta_{N-1}}(\mathbf{p}, t)}{n-1} + \frac{\max\left(0, s - B^N_{\Delta_{N-1}}(\mathbf{p}, t)\Delta_{N-1}\right)}{n-1}. \tag{4.103}$$

As $n \to \infty$, the first additive term $G_{N-1}$ on the right hand side of (4.103) converges to $\gamma_{N-1}$ in mean-square sense due to induction hypothesis. The other two terms converge to 0 in mean-square. The proofs are similar to what showed in (4.88) and (4.89), and we omit the details. Hence, $G_N$ converges to $\gamma_N = \gamma_{N-1}$ in mean square sense:

$$\lim_{n \to \infty} E\left[(G_N - \gamma_N)^2\right] = 0. \tag{4.104}$$

For the case when $\lambda_{\mathbf{p}} > C_N$. Due to Theorem 10, we have

$$\gamma_N = \frac{\sum_{k=1}^{N} \Psi^T \Gamma_N^k \gamma_{k-1} + s}{C_N}. \tag{4.105}$$

Further according to Lemma 11, we have

$$\Delta_N = \frac{Y^N_{\Delta_{N-1}}(\mathbf{p}, t)\Delta_{N-1}}{C_N} + \frac{s}{C_N} + \tilde{I}_N, \tag{4.106}$$

where $t$ is the arrival time of $p_1$ at $L_N$, $\tilde{I}$ is the hop idle time of $L_N$ during the interval $[t, t + \Delta_N]$, and can be expressed as

$$\tilde{I}_N = \max\left(0, B^N_{\Delta_{N-1}}(\mathbf{p}, t)\Delta_{N-1} - s\right). \tag{4.107}$$

Dividing $n - 1$ at both sides of (4.106), we get

$$G_N = \frac{Y_{\Delta_{N-1}}^N(\mathbf{p}, t) G_{N-1}}{C_N} + \frac{s}{(n-1)C_N} + \frac{\tilde{I}_N}{n-1}.$$
(4.108)

The first additive term on the right hand side of (4.108) converges in mean-square to $\lambda_{\mathbf{p}} \gamma_{N-1} / C_N$. We omit the proof details but point out that it requires the condition that the variance of $Y_\delta^N(\mathbf{p}, t)$ is uniformly bounded by some constant for all $\delta$, which we have justified previously. The second term is deterministic, and its square converges to 0 as $n \to \infty$. The third term converges to 0 in mean-square as $n$ increases. To prove this, we first show that $B_{\Delta_{N-1}}^N(\mathbf{p}, t)$ converges in mean-square to 0. Let $P(x)$ be the distribution function of $G_{N-1}$, we have

$$\lim_{n \to \infty} E \left[ \left( B_{\Delta_{N-1}}^N(\mathbf{p}, t) \right)^2 \right]$$
$$= \lim_{n \to \infty} \int_0^\infty E \left[ \left( B_{(n-1)x}^N(\mathbf{p}, t) \right)^2 \right] dP(x)$$
$$= \int_0^\infty \lim_{n \to \infty} E \left[ \left( B_{(n-1)x}^N(\mathbf{p}, t) \right)^2 \right] dP(x),$$
(4.109)

where the interchange between the limit and the integration is valid, because $B_\delta^N(\mathbf{p}, t)$ is distributed in a finite interval $[0, C_N]$ and its second-order moment is uniformly bounded by $C_N^2$ for all $\delta$. Further, recall that for any $x > 0$,

$$\lim_{n \to \infty} E \left[ \left( B_{(n-1)x}^N(\mathbf{p}, t) \right)^2 \right] = 0.$$
(4.110)

Combining (4.109) and (4.110), we get

$$\lim_{n \to \infty} E \left[ \left( B_{\Delta_{N-1}}^N(\mathbf{p}, t) \right)^2 \right] = 0.$$
(4.111)

Using an argument similar to (4.96) and further recalling (4.107), we can easily get

$$\lim_{n \to \infty} E \left[ \left( \frac{\tilde{I}_N}{n-1} \right)^2 \right] = 0.$$
(4.112)

Combining the results for all three additive terms in (4.108), we get the conclusion that when $\lambda_{\mathbf{p}} > C_N$, $G_N$ converges in mean-square to $\lambda_{\mathbf{p}}\gamma_{N-1}/C_N$, which equals to $\gamma_N$ due to (4.101) and Theorem 10. Combining the two cases, we complete the inductive step and the Theorem follows. ∎

### 4.4.4 Discussion

Among the assumptions made in this chapter, some are critical in leading to our results while others are meant to simplify discussions. We point out that the stationarity assumption can be greatly relaxed without harming our major results. However, this comes at the expense of more technical derivations. When cross-traffic arrival is non-stationary, the output dispersion process $G_N(m)$ is no longer an identically distributed random sequence. Consequently, the analysis of probing response curve can not be simply reduced to the analysis of an *individual* output dispersion random variable. Instead, we have to directly examine the sample-path frequency distribution of the output dispersion process $\{G_N(m), 1 \le m \le \infty\}$. We also have to rely on an ASTA assumption on packet-train probing as we did in chapter 3, which we have avoided in the stationary settings of this chapter. We leave as future work the identification of the settings that best fit the Internet traffic environments.

On the other hand, the mean-square ergodicity plays a central role in the proofs for Theorem 13, Theorem 14, and Theorem 15. Although in stochastic process theory, ergodicity is always discussed in the context of stationary process, we can certainly maintain an ergodicity-like condition that is decoupled from stationarity. That is, we can assume (4.24), where $\psi_j$ is just the long-term arrival rate of flow $f_j$. Such a flow, when observed in a large time-

intervals, almost has constant arrival rate. We call this type of cross-traffic flow "asymptotically CRF-like (ACL)". A simple example of non-stationary ACL flow is an on-off cross-traffic flow which alternates between two stationary states. We note that the vast majority of traffic models in stochastic literature are asymptotically CRF-like and consequently that our results have a broad applicability in practice.

Next, we provide experimental evidence for our theoretical results using simulation, testbed experiment, and real Internet measurement data.

## 4.5   Experimental Verification

To get the response curve, we need to obtain the asymptotic average of the probing output dispersions. The period testing and trace-driven testing methods proposed in chapter 3 produce very smooth and accurate curves. Unfortunately, they only work for single-hop paths. In a multi-hop path, we have to rely on measurements using a large number of probing samples. Even though this approach can hardly produce a smooth response curve, the bright side is that it allows us to observe the output dispersion variance, which is reflected by the degree of smoothness of the measured response curve.

### 4.5.1   Testbed Experiments

In our first experiment, we measure in the Emulab testbed [1] the response curves of a three-hop path with the following configuration matrix (all in mb/s)

and one-hop persistent cross-traffic routing

$$\mathbf{H} = \begin{pmatrix} 96 & 96 & 96 \\ 20 & 40 & 60 \end{pmatrix}. \tag{4.113}$$

We generate cross-traffic using three NLANR [2] traces. All the inter-packet delays in each trace are scaled by a common factor so that the average rate during the trace interval becomes the desired value. The trace durations after scaling are 1-2 minutes. We measure the average output dispersions at 100 input rate points, from 1mb/s to 100mb/s with 1mb/s increasing step. For each input rate, we use 500 packet-trains with packet size 1500 bytes. The packet train length $n$ is 65. The inter-probing delay is controlled by an exponentially distributed random variable with sufficiently large mean. The whole experiment lasts for about 73 minutes. All the three traffic traces are made circulatory and they are replayed at random starting points once the previous round is finished. By recycling the same traces in this fashion, we make the cross-traffic last until the experiment ends without creating periodicity. Note that the packet-trains are injected with their input rates so arranged that the 500 trains for each input rate is evenly separated during the whole testing period.

This experiment not only allows us to measure the response curve for $n = 65$, but also for any packet-train length $k$ such that $2 < k < n = 65$, by simply taking the dispersions of the first $k$ packets in each packet-train. Fig. 4.2(a) shows the rate response curve $\tilde{Z}(r_I, s, n)$ for $k = 2, 3, 5, 9, 17, 33$ and 65 respectively. For comparison purposes, we also plot in the figure the multi-hop fluid curve $\tilde{F}(r_I)$ and the single-hop fluid curve $\tilde{S}(r_I)$ of the tight link $L_3$ . The
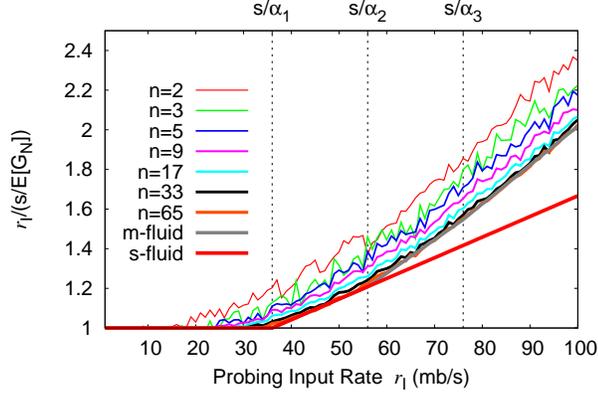
(a) one-hop persistent routing
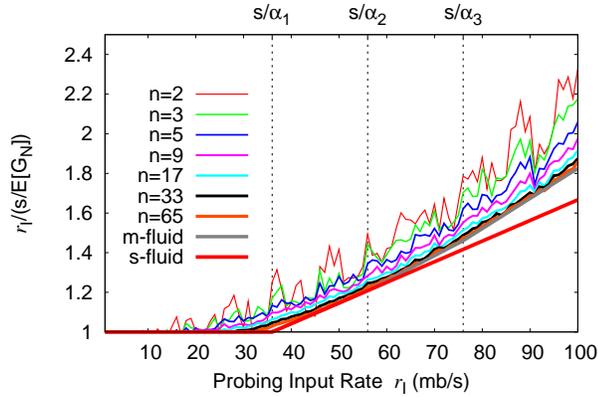


(b) path-persistent routing

Figure 4.2: Measured response curves using different packet train-length in Emulab testbed.

rate response curves $\tilde{Z}(r_I, s, n)$ and $\tilde{S}(r_I)$ are defined as follows

$$\tilde{Z}(r_I, s, n) = \frac{r_I}{s/E[G_N(s/r_I, s, n)]} \qquad (4.114)$$

$$\tilde{S}(r_I) = \max\left(1, \frac{\lambda_b + r_I}{C_b}\right). \qquad (4.115)$$

First note that the multi-hop fluid rate curve comprises four linear segments separated by turning points 36mb/s, 56mb/s, and 76mb/s. The last two linear segments have very close slops and they are not easily distinguishable from

each other in the figure. We also clearly see that the rate curve asymptotically approaches its fluid lower bound as packet-train length $n$ increases. The curve for $n = 65$ almost coincides with the fluid bound. Also note that the smoothness of the measurement curve reflects the variance of the output dispersion random variables. As the packet train length increases, the measured curve becomes smoother, indicating the fact that the variance of the output dispersions is decaying. These observations are all in agreement with those stated in Theorem 15.

Unlike single-hop response curves, which have no deviation from the fluid bound when the input rate $r_I$ is greater than the link capacity, multi-hop response curves usually deviate from its fluid counterpart in the entire input range. As we see from Fig. 4.2(a), even when the input rate is larger than 96mb/s, the measured curves still appear above their multi-hop fluid counterpart. We also see from the figure that the single-hop fluid curve $\tilde{S}$ of the tight link $L_3$ coincides with the multi-hop fluid curve $\tilde{F}$ within the input rate range $(0, 56)$ but falls below $\tilde{F}$ in the input rate range $(56, \infty)$. Consequently, $\tilde{S}$ is only partially obtainable by probing the multi-hop path $\mathcal{P}$ using long packet-trains.

Finally, we explain why we choose the link capacities to be 96mb/s instead of the fast ethernet capacity 100mb/s. In fact, we did set the link capacity to be 100mb/s. However, we noticed that the measured curves can not get arbitrarily close to their fluid bound $\tilde{F}$ computed based on the fast ethernet capacity. Using pathload to examine the true capacity of each Emulab link, we found that their IP layer capacities are in fact 96mb/s, not the same as their nominal value 100mb/s.

In our second experiment, we change the cross-traffic routing to path-persistent while keeping the path configuration matrix the same as given by

(4.113). That is,

$$
\mathbf{R}^T \Psi = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^T \times \begin{pmatrix} 20 \\ 20 \\ 20 \end{pmatrix} = \begin{pmatrix} 20 \\ 40 \\ 60 \end{pmatrix}. \tag{4.116}
$$

We repeat the same packet-train probing experiment and the results are plotted in Fig. 4.2(b). The multi-hop fluid rate curve $\tilde{F}$ still coincides with $\tilde{S}$ in the input rate range $(0, 56)$. When input rate is larger than 56mb/s, $\tilde{F}$ positively deviates from $\tilde{S}$. However, the amount of deviation is smaller than that in one-hop persistent routing. The measured curves approach to the fluid lower bound $\tilde{F}$ with decaying variance as packet-train length increases. For $n = 65$, the measured curve becomes hardly distinguishable from $\tilde{F}$.

We have also conducted experiment with paths of more hops and much complicated routing pattern and path configuration. Results obtained (not shown for brevity) all perfectly support our theory. Next, we examine the impact of probing packet size. Since in practice, packet size is usually limited by ethernet MTU and can not be more than 1500 bytes. We decide to use ns2 simulation, where packet size can be set to any large value we wish.

## 4.5.2   Simulation Results

The path settings and cross-traffic used in our simulation are the same as those in Emulab testbed experiments. However, the link capacities in ns2 simulation are what they are set to be – 100mb/s. In the first simulation experiment, cross-traffic routing is one-hop persistent. We use packet-pairs of different sizes to measure the rate response curves. For each probing packet size, we probe the path at 45 input rates, from 10 mb/s to 100 mb/s with 2mb/s increasing step.

For each input rate, we use 500 packet-pairs to estimate the average output rate $s/E[G_N]$. Fig. 4.3(a) plots the rate curves for probing packet sizes 500, 1000, 2000, 4000, and 8000 (all in bytes). We see that as packet-size increases, the response curve approaches its multi-hop fluid counterpart. This trend is obvious even though with the largest size used (8,000bytes), the convergence is still not sufficient in certain input rate range.

In the second simulation experiment, we change the cross-traffic routing to path-persistent while keep all other factors the same. The rate curves associated with the five different probing packet sizes are plotted in Fig. 4.3(b), where we see the same convergence pattern even though the multi-hop fluid curve becomes different.

### 4.5.3 Real Internet Measurements

We conducted packet-train probing experiments on several Internet paths in RON testbed to verify our analysis on real networks. Since neither the path configuration nor the cross-traffic routing information is available for those Internet paths, we are unable to provide the fluid bounds. Therefore, we verify our theory by observing the convergence of the measured curves to a piece-wise linear curve as packet-train length increases.

In the first experiment, we measure the rate response curve of the path from the RON node **lulea** at Sweden to the RON node at CMU. The path has 19 hops and a fast-ethernet bottleneck, as we find out using traceroute and pathrate. We probe the path at 29 different input rates, from 10mb/s to 150mb/s with a 5mb/s increasing step. For each input rate, we use 200 packet-trains of 33-packet length to estimate the output probing rate $s/E[G_N]$. The whole experiment
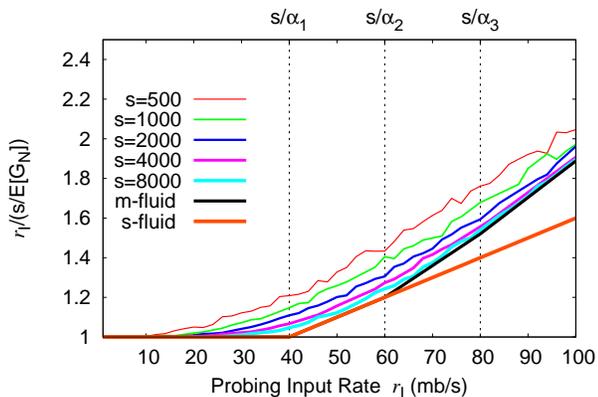
(a) one-hop persistent routing
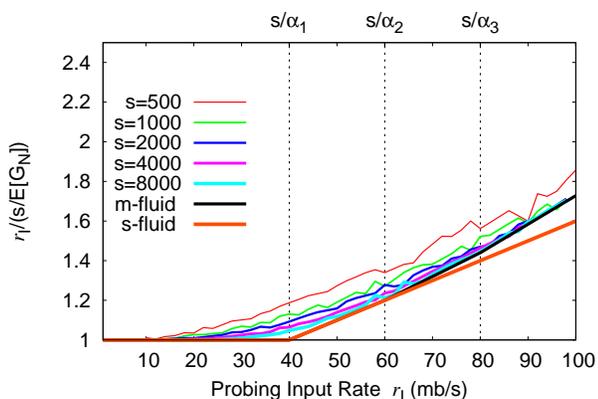


(b) path persistent routing

Figure 4.3: Measured response curves using different packet sizes in ns2 simulation.

took about 24 minutes. Again, the 200 packet-trains for each of the 29 input rates are so arranged that they are approximately evenly separated during the 24-minute testing period. The measured rate response curves associated with packet-train length 2, 3, 5, 9, 17, and 33 are plotted in Fig. 4.4(a), where we see that the response curve approaches a piece-wise linear bound as packet-train length increases. At the same time, response curves measured using long trains are smoother than those measured using short trains, indicating the decaying

variance of output dispersions. In this experiment, the curve measured using probing trains of 33-packet length exhibits sufficient smoothness and clear piecewise linearity. We only observed two linear segments from the figure, possibly because the other linear segments are located outside the input rate range we measured.

Based on (4.20), we apply linear regression on the second linear segment to compute the capacity $C_b$ and the cross-traffic intensity $\lambda_b$ at the bottleneck link and get $C_b = 96$mb/s and $\lambda_b = 2$mb/s. Using these results, we retroactively plot the single-hop fluid bounds and observe that it almost overlaps with the measured curve using packet-trains of 33-packet length. Notice that the bottleneck link is under very light utilization during our 24-minute measurement period. We can infer based on our measurement that the available bandwidth of the path is constrained mainly by the capacity of the bottleneck link and that the probing packet-trains have undergone significant interaction with cross-traffic at non-bottleneck links. Otherwise, according to Theorem 3, the response curves measured using short train lengths would not have appeared above the single-hop fluid bound when input rate is larger than the tight link capacity 96mb/s. We believe that the tight link of the path is one of the last-mile lightly utilized fast-ethernet links and that the backbone links are transmitting significant amount of cross-traffic even though they still have available bandwidth much more than the fast-ethernet capacity. Also notice that similar to our testbed experiments, fast-ethernet links only have 96mb/s IP-layer capacity.

We repeat the same experiment on another path from RON node **pwh** at Sunnyvale California to NYU RON node. This path has 13 hops and fast-ethernet bottleneck capacity. Due to substantial cross-traffic burstiness along the path, we use packet-trains of 129-packet length in our probing experiment.
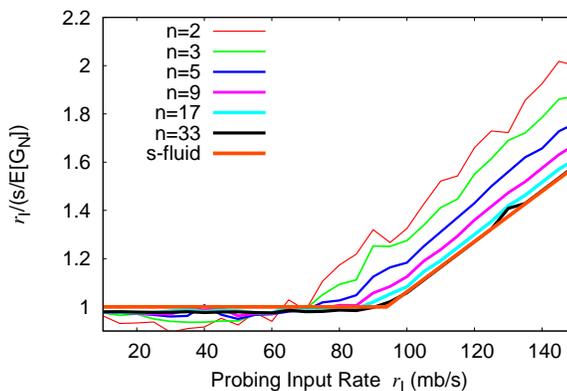
(a) lulea $\rightarrow$ CMU
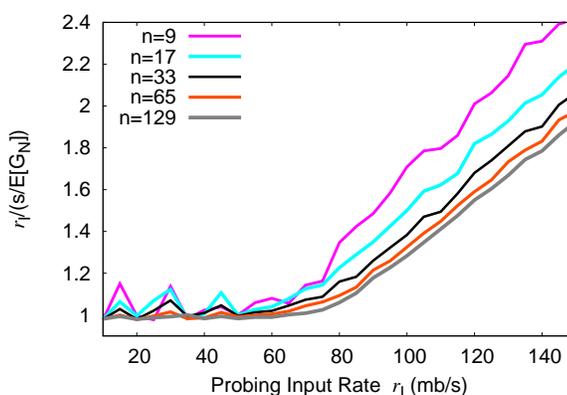


(b) pwh $\rightarrow$ NYU

Figure 4.4: Measured response curves of two Internet paths in RON testbed .

The other parameters such as the input rates and the number of trains used for each rate are all the same as in the previous experiment. The whole measurement duration is about 20 minutes. The measured response curves are plotted in Fig. 4.4(b). As we see, the results exhibit more measurement variability compared to the lulea$\rightarrow$CMU path. However, as packet-train length increases, the variability is gradually smoothed out and the response curve converges to a piece-wise linear bound. We again apply linear regression on the response curve for 129 train length to obtain tight link information. We get $C_b = 80$mb/s

and $\lambda_b = 3$mb/s, which does not agree with pathrate. We believe that it is because there are links along the path with very similar available bandwidth. Consequently, the second linear segment become too short to detect. The linear segment we are acting linear regression upon is probably a later one, whose congestible hop set includes several links. This experiment confirms our analysis, at the same time shows some of the potential difficulties in exacting tight link information from the response curves.

## 4.6 Implications

We now discuss the implications of our results on existing measurement techniques. Except for pathChirp, all other techniques such as TOPP, pathload, PTR, and Spruce are related to our analysis.

### 4.6.1 TOPP

TOPP is based on multi-hop fluid rate response curve $\tilde{F}$ with one-hop persistent cross-traffic routing. TOPP uses packet-pairs to measure the real rate response curve $\tilde{Z}$, and assumes that the measured curve will be the same as $\tilde{F}$ when large number of packet-pairs are used. However, our analysis shows that the real curve $\tilde{Z}$ is different from $\tilde{F}$, especially when packet-trains of short length are used (e.g., packet-pairs). Note that there is not much path information in $\tilde{Z}$ that is readily extractable unless it is sufficiently close to its fluid counterpart $\tilde{F}$. Hence, to put TOPP to work in practice, one must use long packet-trains instead of packet-pairs. We also point out that when the fluid curve $\tilde{F}$ is obtained with decent precision, we might be able to extract from $\tilde{F}$ the capacity information $C_b$ and the cross-traffic intensity information $\lambda_b$ of the tight link. However,
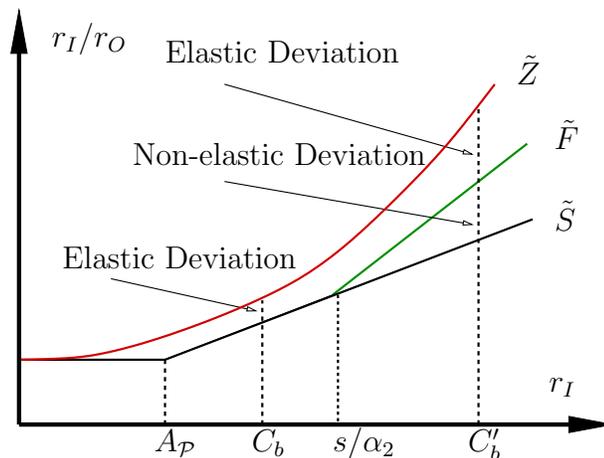
Figure 4.5: Illustration of two types of curve deviations.

it is unlikely to extract such information for non-bottleneck links due to their dependence on cross-traffic routing and practical difficulties in linear segment detection.

## 4.6.2 Spruce

Using our notations system , we can write spruce's available bandwidth estimator as follows

$$C_b \left( 1 - \frac{G_N(s/C_b, s, n) - s/C_b}{s/C_b} \right),$$

(4.117)

where the probing packet size $s$ is set to 1500bytes, the packet-train length $n = 2$, and the bottleneck link capacity $C_b$ is assumed known.

It is showed in chapter 3 that spruce estimator is unbiased in single-hop paths regardless of the packet-train parameters $s$ and $n$. This means that the statistical mean of (4.117) is equal to $A_{\mathcal{P}}$ for any $s > 0$ and any $n \geq 2$. Next, we derive a necessary condition for this unbiasedness property to hold in a

multi-hop path $\mathcal{P}$. Note that

$$E\left[C_b\left(1 - \frac{G_N(s/C_b, s, n) - s/C_b}{s/C_b}\right)\right] = C_b\left(2 - \tilde{Z}(C_b, s, n)\right), \qquad (4.118)$$

where $\tilde{Z}(C_b, s, n)$ is the rate response of path $\mathcal{P}$ at input rate $C_b$ with packet-train parameters $s$ and $n$. Unbiasedness property of spruce estimator is satisfied only when

$$C_b(2 - \tilde{Z}(C_b, s, n)) = A_{\mathcal{P}} = C_b - \lambda_b, \qquad (4.119)$$

which is equivalent to the following condition

$$\tilde{Z}(C_b, s, n) = \frac{\lambda_b + C_b}{C_b} = \tilde{S}(C_b). \qquad (4.120)$$

This means that at the input rate point $C_b$, the real rate response of path $\mathcal{P}$ must be equal to the single-hop fluid rate response at the tight link of $\mathcal{P}$.

This condition is usually not satisfied. Instead, due to Theorem 11 and Property 4, we have

$$\tilde{Z}(C_b, s, n) \geq \tilde{F}(C_b) \geq \tilde{S}(C_b). \qquad (4.121)$$

Combining (4.121) and (4.118), we see that (4.117) is a negatively biased estimator of $A_{\mathcal{P}}$. The amount of bias is given by

$$\begin{aligned} C_b&\left(\tilde{Z}(C_b, s, n) - \tilde{S}(C_b)\right) \qquad\qquad\qquad (4.122)\\ &= C_b\left(\tilde{Z}(C_b, s, n) - \tilde{F}(C_b)\right) + C_b\left(\tilde{F}(C_b) - \tilde{S}(C_b)\right). \end{aligned}$$

The first additive term in (4.122) is the measurement bias caused by the curve deviation of $\tilde{Z}$ from $\tilde{F}$ at input rate $C_b$, which vanishes as $n \to \infty$ due to Theorem 15. Hence we call it "elastic bias". The second additive term is the portion of measurement bias caused by the curve deviation of $\tilde{F}$ from $\tilde{S}$ at input rate $C_b$, which remains constant with respect to packet-train parameters

$s$ and $n$. Therefore it is "non-elastic bias". We illustrate the two types of curve deviations in Fig. 4.5. Note that when $C_b < s/\alpha_2$, non-elastic bias is 0. Further recall that $s/\alpha_2 \geq A_{b2}$ as stated in Property 3. Hence, a sufficient condition for zero "non-elastic" bias is $C_b \leq A_{b2}$. Conceptually, elastic deviation stems from cross-traffic burstiness and non-elastic deviation is a consequence of "multi-hop" effect.

In Table 4.2, we give the amount measurement bias caused by the two types of curve deviations in both the Emulab testbed experiments and real Internet probing measurements on the path from lulea to CMU. Note that in the testbed experiment using a 3-hop path with one-hop persistent routing, spruce suffers about 74mb/s measurement bias, which is twice as much as the actual path available bandwidth 36mb/s. In the second Emulab experiment using path-persistent cross-traffic, the measurement bias is drastically reduced to 38.8mb/s, which however is still more than the actual available bandwidth. In both cases, spruce estimator converges to negative values. We use spruce to estimate the two paths and it does give 0mb/s results in both cases. For the Internet path from lulea to CMU, spruce suffers 24mb/s negative bias and produces a measurement result less than 70mb/s, while the real value is around 94mb/s. We also use pathload to measure the three paths and it produces pretty accurate results.

The way to reduce elastic-bias is to use long packet-trains instead of packet-pairs. In the lulea→CMU experiment, using packet-trains of 33-packet, spruce can almost completely overcome the 24mb/s bias and produce an accurate result. However, there are two problems of using long packet-trains. First, there is not a deterministic train length that guarantees negligible measurement bias on any network path. Second, when router buffer space is limited and packet-train length are too large, the later probing packets in each train may experience

| experiment | elastic bias | non-elastic bias | total bias |
|:---:|:---:|:---:|:---:|
| Emulab-1 | $0.56 \times 96$ | $0.315 \times 96$ | 74.4 |
| Emulab-2 | $0.28 \times 96$ | $0.125 \times 96$ | 38.8 |
| lulea-cmu | $0.25 \times 96$ | 0 | 24 |

Table 4.2: Spurce bias in Emulab and Internet experiment (in mb/s).

frequent loss, making it impossible to accurately measure $\tilde{F}(C_b)$. After all, spruce uses input rate $C_b$, which can be too high for the bottleneck router to accommodate long packet-trains. On the other hand, note that non-elastic bias is an inherit problem for spruce. There is no way to overcome it by adjusting packet-train parameters.

### 4.6.3   PTR and pathload

PTR searches the first "turning point" in the response curve $\tilde{Z}(r_I, s, n)$ and takes the input rate of the turning point as the path available bandwidth $A_{\mathcal{P}}$. This method can produce accurate result when the real response curve $\tilde{Z}$ is close to $\tilde{F}$, which requires packet-train length $n$ be sufficiently large. Otherwise, PTR is also negative biased and produces underestimation of $A_{\mathcal{P}}$. The minimum packet-train length needed is dependent on the path conditions. The current version of PTR use packet train length $n = 60$, which can produce pretty accurate results for the paths experimented in this chapter.

Pathload is in spirit similar to PTR. However, it searches the available bandwidth region by detecting the one-way-delay increasing trend within a packet-train, which is different from examining wether the rate response $\tilde{Z}(r_I, s, n)$ is greater than one [19]. However, since there is a strong statistical correlation

between a high rate response $\tilde{Z}(r_I, s, n)$ and the one-way-delay increasing tend within packet-trains, our analysis can explain the behavior of pathload to a certain extent. It is reported in [18] that pathload underestimates available bandwidth when there are multiple tight links along the path. This is because in the input rate range $(0, A_{\mathcal{P}})$, the deviation of $\tilde{Z}(r_I, s, n)$ from $\tilde{F}$ is maximized when non-bottleneck links have the same available bandwidth as $A_{\mathcal{P}}$, given that the other factors are kept the same. It is our new observation that by further increasing the packet-train length, the underestimation can be mitigated in pathload.

Our analysis sheds new light on the essence of available bandwidth. Even through multiple tight links cause one-way-delay increasing trend for packet-trains with input rate less than $A_{\mathcal{P}}$, this is *not* an indication that the network can not sustain such an input rate. Rather, the increasing trend is a *transient* phenomenon resulting from "probing intrusion residual", and it disappears when the input packet-train is sufficiently long. The concept of path available bandwidth has an intrinsic nature regardless of path configuration. It is the largest *long-term* input rate the network can sustain without causing steady one-way-delay increasing.

## 4.7   Conclusion

In this chapter, we provide a theoretical understanding of packet-train bandwidth estimation in a multi-hop path with arbitrarily routed cross-traffic flows. Our main contributions includes the derivation for both the multi-hop fluid response curve and the real response curve, an investigation of the convergence properties of the real response curve with respect to packet-train parameters,

and the implications of our analysis outcomes on existing techniques.

We leave as future work the investigation of new approaches that help detect and eliminate the measurement bias caused by bursty cross-traffic in multi-hop paths.

# Chapter 5

# Summary

In this chapter, we briefly summarize the main results of this thesis and present several potential applications of our analysis. We then point out our future research directions.

## 5.1   Main Results

Among all results we obtained in this dissertation, Theorem 10, Theorem 4, and Theorem 15 are of the most practical relevance. Existing measurement techniques are mostly based on the single-hop fluid response curve. The validity of this foundation is that in certain input probing range, the single-hop fluid response curve is a lower bound of the real response curve, *approachable* when packet-train length is sufficiently large. On the other hand, the inadequacy of existing techniques comes from the lack of understanding for the significance of both the packet-train length and the input probing rate.

## 5.2   Practical Applications

Our analysis outcomes have several important practical applications, which we briefly discuss in the following.

The first application of our theory is that it suggests a well-grounded methodology to measure several characteristics of the available bandwidth bottleneck link such as its capacity and utilization. Previous work either measure path available bandwidth without any knowledge of the bottleneck capacity, or assume it is the same as the narrow link capacity, which can be measured using capacity estimation tools. The capacity information about the bottleneck link is very useful. It tells weather the path available bandwidth is mainly constrained by the link capacity (when utilization is low) or it is constrained by the heavy utilization. This is especially useful for network managers to conduct effective capacity planning.

Being able to measure tight link capacity also allows us to verify the previous assumption that the tight link is also the narrow link. Several techniques, such as IGI [15], Spruce [38], and Delphi [36] rely on this assumption to conduct measurements. Therefore, a verification of this assumption is important in that it determines the applicability of these tools in real Internet environments.

A second application of our multi-hop theory is to locate the bottleneck link using packet-train probing. By collecting the output dispersion random variable $G_i(g_I, s, n)$ at each link[1], we can compute the available bandwidth for each path prefix, consequently locating the link that constrains path available bandwidth. There are several proposals for bottleneck localization [14], [43], however, they are all based on fluid analysis. Our multi-hop theory can help

---

[1]This can be done using the approach in [14].

improve the accuracy of current bottleneck localization tools.

## 5.3   Future Work

There are several research directions we are interested in pursuing in our future work.

In the single-hop analysis, we made two stability assumptions on cross-traffic arrival and most of our results rely on PASTA sampling. In the multi-hop analysis, we assumed stationary increments on cross-traffic arrival, consequently the output dispersion process $\{G_N(m)\}$ has time-invariant distribution. The results obtained therein requires no conditions on inter-probing pattern. The cross-traffic assumptions made in multi-hop analysis is stronger than the ones in single-hop analysis. It still remains as an open problem which assumptions are best suited for Internet cross-traffic environment. Understanding the statistical structure of the output dispersion process in the current Internet is important in the design of inter-probing pattern in bandwidth measurement techniques. We are interested in exploring the answer to this problem through extensive real measurements in a near future.

In the single-hop analysis, we proposed a method called "trace-driven testing" that can compute the single-hop response curve with high accuracy. The computed curves are smooth and monotonic. In multi-hop analysis, we relied on experimental measurement of the real response curves. The measured curves exhibit substantial variations and less accuracy. We are interested in finding new ways of computing the multi-hop response curves with more accuracy and smoothness, especially for short probing packet-trains.

We proved that the response curves approach the fluid bound as packet-

train length increases. An interesting problem is to investigate the speed of this convergence and its deciding factors. In the single-hop analysis, we find in experiments that the convergence can be accurately modeled using a power-law function of the packet-train length. We are interested in extending this result to multi-hop path and find better justification for this phenomenon. Understanding the convergence pattern allows us to infer the fluid bound based on several response curves associated with short packet-train lengths and to avoid using excessively long packet-trains to overcome measurement bias.

# Appendix A

# Remarks on Cross-Traffic Stationarity

Cross-traffic is stationary if the cumulative traffic arrival process $\{V(t)\}$ has stationary increments, which also implies that the $\delta$-interval cross-traffic intensity process $\{Y_\delta(t)\}$ is a stationary process for all $\delta > 0$. Assuming ergodicity, stationary traffic arrival leads to the intensity stability assumption made chapter 3. It is also well established that stationary traffic arrival, when its long term rate $\lambda$ is less than the hop capacity $C$, leads to hop workload stability [29]. Hence, the results in chapter 3 are applicable to stationary cross-traffic. In our experiment, the two cross-traffic PCS and PUS are stationary cross-traffic.

Note that, however, a lot of traffic types that are suited for stochastic modeling are non-stationary. *On/off* traffic is one such example, which belongs to regenerative traffic and often is at most asymptotically stationary. More examples include time dependent Poisson traffic, transition-modulated traffic, and even most of the renewal traffic. By avoiding stationarity assumption, our conclusions are applicable to virtually arbitrary cross-traffic that can be stochas-

tically modeled. The two stability assumptions are also arguably the weakest conditions of cross-traffic measurability.

# Appendix B

# Workload Stability of the Four Traces

We show that the four traffic traces used in the chapter 3 all lead to hop workload stability. We omit CBR due to its triviality. For PCS, PUS, and POF, we first apply queuing theory to calculate their hop workload time averages. We then prove that the existence of workload time average implies workload stability.

We use $\gamma$ to denote the average cross-traffic arrival rate in packet per second, $\mathbf{d}_n$ to denote the packet-delay sample-path, $\mathbf{S}_n$ to denote the packet service time sample-path. The following is a basic result in queueing theory [42, pages 279]:

$$E[W(t)] = \gamma E[\mathbf{S}_n]E[\mathbf{d}_n] + \gamma E[\mathbf{S}_n^2]/2. \tag{B.1}$$

We now apply (B.1) to calculate the workload sample-path time-average for PCS, PUS, and POF. First note that $\gamma = 500$ packets/sec for all three traffic traces.

In PCS, since packet size is constantly 750 bytes, the sample-path mean of packet service time is $E[\mathbf{S}_n] = 6 \times 10^{-4}$s and $E[S_n^2] = 3.6 \times 10^{-7}s^2$. Further

note that due to PASTA, $E[W(t)] = E[\mathbf{d}_n]$. Hence, we have:

$$E[W(t)] = 500 \times 6 \times 10^{-4} \times E[W(t)] + 500 \times 3.6 \times 10^{-7}/2. \qquad \text{(B.2)}$$

Compute $E[W(t)]$ from (B.2), we get $E[W(t)] = 128.57\mu$ s.

In PUS, since packet size is uniformly distributed in $[1, 1500]$ bytes, the sample-path mean of packet service time is $E[\mathbf{S}_n] = 6 \times 10^{-4}$s. The second moment of packet service time is $E[\mathbf{S}_n^2] = 4.8 \times 10^{-7}s^2$. Further note that due to PASTA, $E[W(t)] = E[\mathbf{d}_n]$. Hence, we have:

$$E[W(t)] = 500 \times 0.0006 \times E[W(t)] + 500 \times 4.8 \times 10^{-7}/2. \qquad \text{(B.3)}$$

Compute $E[W(t)]$ from (B.3), we get $E[W(t)] = 171.43\mu$ s.

In POF, since packet size is constantly 750 bytes, the sample-path mean of packet service time is $E[\mathbf{S}_n] = 6 \times 10^{-4}$s. The second moment of packet service time is $E[\mathbf{S}_n^2] = 3.6 \times 10^{-7}s^2$. Further note that in POF, all packets come see empty queue. Thus, $E[\mathbf{d}_n] = 0$ and we have:

$$E[W(t)] = 500 \times 3.6 \times 10^{-7}/2 = 90\mu s. \qquad \text{(B.4)}$$

We plot the average workload function $\mathcal{W}(t) = \int_0^t W(u)du/t$ for the three cross-traffic traces in Figure B.1. It is clear that the plot agrees with queuing theoretic computation.

**Theorem 16** *If* $\lim_{t\to\infty} \mathcal{W}(t)$ *exists and is finite, then* $\exists t_0$, *for* $\forall t > t_0$, $W(t) < \sqrt{t}$.

   **Proof:**   Let

$$\lim_{t\to\infty} \mathcal{W}(t) = \lim_{t\to\infty} \frac{\int_0^t W(u)du}{t} = k. \qquad \text{(B.5)}$$
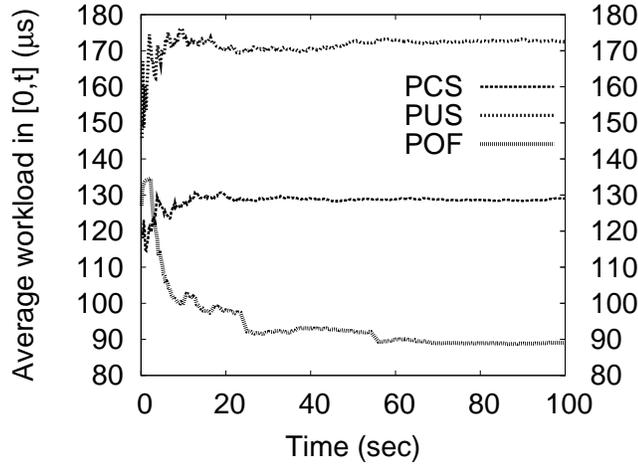
Figure B.1: Average hop workload $\mathcal{W}(t)$ for PCS, PUS, and POF.

Suppose theorem 16 does not hold, then there exists an infinite series $\{t_n\}$, such that $\lim_{n\to\infty} t_n = \infty$ and $W(t_n) \geq \sqrt{t_n}$ for $\forall n$. Due to basic real analysis theorem,

$$\lim_{n\to\infty} \frac{\int_0^{t_n+\sqrt{t_n}} W(u)du}{t_n + \sqrt{t_n}} = k. \tag{B.6}$$

However, due to the sample-path properties of $W(t)$,

$$\int_0^{t_n+\sqrt{t_n}} W(u)du \geq \int_0^{t_n} W(u)du + \frac{t_n}{2}. \tag{B.7}$$

Thus, we have

$$\begin{aligned}
\lim_{n\to\infty} & \frac{\int_0^{t_n+\sqrt{t_n}} W(u)du}{t_n + \sqrt{t_n}} \\
&\geq \lim_{n\to\infty} \left( \frac{\int_0^{t_n} W(u)du}{t_n + \sqrt{t_n}} + \frac{t_n}{2(t_n + \sqrt{t_n})} \right) \\
&= k + \frac{1}{2}.
\end{aligned} \tag{B.8}$$

The contradiction proves this theorem. ∎

Theorem 16 shows that when workload sample-path has a finite limiting time average, then it is asymptotically bounded by $\sqrt{t}$. This immediately leads

to the following:

$$\lim_{t \to \infty} \frac{W(t)}{t} = 0. \tag{B.9}$$

# Bibliography

[1] Emulab. http://www.emulab.net.

[2] National Laboratory for Applied Network Research. http://www.nlanr.net.

[3] M. Abramowitzand and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 1972.

[4] K. G. Anagnostakis, M. B. Greenwald, and R. S. Ryger. cing: Measuring network-internal delays using only existing infrastructure. In *IEEE INFOCOM*, April 2003.

[5] H. Balakrishnan, M. Stemm, S. Seshan, and R. H. Katz. Analyzing stability in wide-area network performance. In *ACM SIGMETRICS*, June.

[6] S. Banerjee and A. Agrawala. Estimating available capacity of a network connection. In *In Proceedings IEEE International Conference on Networks*, Sept. 2001.

[7] J. Bolot. Characterizing end-to-end packet delay and loss in the internet. In *ACM SIGCOMM*, 1993.

[8] R. Carter and M. Crovella. Measuring bottleneck link speed in packet-switched networks. *Internation Journal on Performance Evaluation*, 2728, 1996.

[9] C. Dovrolis, P. Ramanathan, and D. Moore. What do packet dispersion techniques measure? In *IEEE INFOCOM*, April 2001.

[10] C. Dovrolis, P. Ramanathan, and D. Moore. Packet dispersion techniques and a capacity estimation methodology. In *IEEE/ACM Transaction on Networking*, March 2004.

[11] A. B. Downey. Using pathchar to estimate internet link characteristics. In *ACM SIGCOMM*, 1999.

[12] K. Harfoush, A. Bestavros, and J. Byers. Measuring bottleneck bandwidth of targeted path segments. In *INFOCOM*, March 2003.

[13] G. He and J. Hou. On exploiting long range dependence of network traffic in measuring cross traffic on an end-to-end basis. In *IEEE INFOCOM*, March 2003.

[14] N. Hu, L. E. Li, Z. M. Mao, P. Steenkiste, and J. Wang. Locating internet bottlenecks: Algorithms, measurements, and implications. In *ACM SIGCOMM*, August 2004.

[15] N. Hu and P. Steenkiste. Evaluation and characterization of available bandwidth probing techniques. *IEEE JSAC Special Issue in Internet and WWW Measurement, Mapping, and Modeling*, 3rd Quarter 2003.

[16] V. Jacobson. Congestion avoidance and control. In *ACM SIGCOMM*, 1988.

[17] V. Jacobson. pathchar – a tool to infer characteristics of internet paths. In *Slides available from ftp://ftp.ee.bl.gov/pathchar/msri-talk.pdf*, 1997.

[18] M. Jain and C. Dovrolis. End-to-end available bandwidth: measurement methodology, dynamics, and relation with tcp throughput. In *ACM SIGCOMM*, August 2002.

[19] M. Jain and C. Dovrolis. Ten fallacies and pitfalls in end-to-end available bandwidth estimation. In *ACM IMC*, October 2004.

[20] S. Keshav. A control-theoretic approach to flow control. In *ACM SIGCOMM*, 1991.

[21] K. Lai and M. Baker. Measuring bandwidth. In *IEEE INFOCOM*, 1999.

[22] K. Lai and M. Baker. Measuring link bandwidths using a deterministic model of packet delay. In *ACM SIGCOMM*, 2000.

[23] X. Liu, K. Ravindran, B. Liu, and D. Loguinov. Single-hop probing asymptotics in available bandwidth estimation: Sample-path analysis. In *ACM IMC*, October 2004.

[24] W. Matragi, K. Sohraby, and C. Bisdikian. Jitter calculus in atm networks: Multiple nodes. *IEEE/ACM Transctions on Networking*, 5(1):122–133, 1997.

[25] B. Melamed and D. Yao. The asta property. *Advances in Queueing: Theory, Methods and Open Problems*, pages 195–224, 1995.

[26] B. Melander, M. Bjorkman, and P. Gunningberg. A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In *IEEE Globecom Global Internet Symposium*, November 2000.

[27] B. Melander, M. Bjorkman, and P. Gunningberg. First-come-first-served packet dispersion and implications for tcp. In *IEEE GLOBECOM*, 2002.

[28] B. Melander, M. Bjorkman, and P. Gunningberg. Regression-based available bandwidth measurements. In *SPECTS*, July 2002.

[29] E. Muhammad and S. Stidham. *Sample-path analysis of queueing systems.* Kluwer Academic Publishers, 1999.

[30] Y. Ohba, M. Murata, and H. Miyahara. Analysis of interdeparture processes for bursty traffic in atm networks. *IEEE Journal on Selected Areas in Communications*, 9, 1991.

[31] A. Pasztor and D. Veitch. Active probing using packet quartets. In *ACM SIGCOMM IMW*, 2002.

[32] A. Pasztor and D. Veitch. On the scope of end-to-end probing methods. *IEEE Communication Letters*, 2002.

[33] A. Pasztor and D. Veitch. The packet size dependence of packet pair like methods. In *IWQoS02*, 2002.

[34] V. Paxson. End-to-end Internet packet dynamics. *IEEE/ACM Transactions on Networking*, 7(3):277–292, 1999.

[35] R. S. Prasad, C. Dovrolis, and B. A. Mah. The effect of layer-2 store-and-forward devices on per-hop capacity estimation. In *IEEE INFOCOM*, 2003.

[36] V. Ribeiro, M. Coates, R. Riedi, S. Sarvotham, B. Hendricks, and R. Baraniuk. Multifractal cross traffic estimation. In *Proc. of ITC Specialist Seminar on IP Traffic Measurement*, September 2000.

[37] V. Ribeiro, R. Riedi, R. Baraniuk, J. Navratil, and L. Cottrell. pathchirp: Efficient available bandwidth estimation for network paths. In *Passive and Active Measurement Workshop*, 2003.

[38] J. Strauss, D. Katabi, and F. Kaashoek. A measurement study of available bandwidth estimation tools. In *ACM IMC*, 2003.

[39] W. Szczotka. Stationary representation of queues. i. *Advance in Applied Probability*, 18:815–848, 1986.

[40] W. Szczotka. Stationary representation of queues. ii. *Advance in Applied Probability*, 18:849–859, 1986.

[41] R. Wolff. Poisson arrivals see time averages. *Operations Research*, 30(2):223–231, 1982.

[42] R. Wolff. *Stochastic modeling and the theory of queues.* Prentice hall, 1989.

[43] D. Zhang, W. Huang, and C. lin. Locating the tightest link of a network path. In *ACM SIGMETRICS*, June 2004.

[44] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker. On the constancy of internet path properties. In *ACM SIGCOMM Internet Measurement Workshop*, November 2001.