Queueing Syst (2009) 63: 59–107
DOI 10.1007/s11134-009-9150-9

INVITED PAPER

Inverse problems in queueing theory and Internet probing

F. Baccelli · B. Kauffmann · D. Veitch

Received: 18 May 2009 / Revised: 29 October 2009 / Published online: 26 November 2009

Abstract Queueing theory is typically concerned with the solution of direct problems, where the trajectory of the queueing system, and laws thereof, are derived based on a complete specification of the system, its inputs and initial conditions. In this paper we point out the importance of inverse problems in queueing theory, which aim to deduce unknown parameters of the system based on partially observed trajectories. We focus on the class of problems stemming from probing based methods for packet switched telecommunications networks, which have become a central tool in the measurement of the structure and performance of the Internet. We provide a general definition of the inverse problems in this class and map out the key variants: the analytical methods, the statistical methods and the design of experiments. We also contribute to the theory in each of these subdomains. Accordingly, a particular inverse problem based on product-form queueing network theory is tackled in detail, and a number of other examples are given. We also show how this inverse problem viewpoint translates to the design of concrete Internet probing applications.

Keywords Queueing theory · Inverse problem · Internet probing

The work of F. Baccelli and B. Kauffmann is funded in part by the Euro NF Network of Excellence and by the French ANR Cmon project.

F. Baccelli (✉) · B. Kauffmann
INRIA, École Normale Supérieure, Équipe TREC, 23 avenue d'Italie, CS 81321,
75214 Paris Cedex 13, France
e-mail: francois.baccelli@ens.fr

B. Kauffmann
e-mail: bruno.kauffmann@ens.fr

D. Veitch
ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN), CUBIN is an affiliated program of National ICT Australia (NICTA), Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Australia
e-mail: dveitch@unimelb.edu.au

Mathematics Subject Classification (2000) 60G10 · 60G12 · 60K25 · 62F10 · 62F12 · 62K05 · 62M05 · 62M09 · 68M20 · 68Q87 · 90B15 · 90B18 · 90B22

1 Introduction

Consider a dynamical system governed by some known evolution equation. In a *direct* problem, the parameters of this dynamical system are known and the goal is to calculate the associated 'trajectory'. In an *inverse* problem, one or more trajectories are observed and, using the evolution equations of the system, one tries to deduce (some of) the parameters which gave rise to those trajectories.

A typical example of an inverse problem is in acoustics. In the direct problem, the parameters could be the location and shape of some obstacle as well as some input signals with a given spatial and temporal structure. These parameters, when used together with the theory of wave propagation and scattering, allow one to determine the acoustic signal at any location and time. A classical inverse problem consists in selecting appropriate input signals, measuring the resulting acoustic signal at certain locations where such measurements are possible, and then leveraging the shape of the solution of the direct problem to determine the unknown location and shape of the obstacle.

Inverse problems are in fact ubiquitous in physics, and have well established incarnations in many other fields such as fluid dynamics and electromagnetism. They have major applications in seismology, geophysics, medical imaging, and industrial process monitoring, to quote just a few.

The following toy example exhibits many of the key features of inverse problems which we consider in this paper, and allows us to introduce some terminology.

A mass initially at height y_0 and with vertical speed v_0 has a trajectory given by the *direct equation* $y(t) = y_0 + v_0 t - g\frac{t^2}{2}$. Assume that the initial conditions y_0 and v_0 are hidden to some observer, who can only glimpse the trajectory at n different epochs, and assume each glimpse allows the observer to make an accurate measurement, an *observation*, of the mass's location. Our inverse problem consists in determining the *unknown parameters* y_0 and v_0 from the observations. It is easy to see that if $n \geq 2$ and if the observer knows this direct evolution equation, then the observations suffice for him to determine the unknown parameters unambiguously. If $n = 1$, the observer can only infer a linear relationship between y_0 and v_0, and the inverse problem is *ill-posed* or ambiguous, lacking a unique solution. Furthermore, if g is unknown then the triple (y_0, v_0, g) can also be determined from such observations, and this is in fact one of the ways for estimating local values of g.

Note that the direct equation of our toy system lives in continuous time (and space). A natural inverse problem is to determine the parameters given observations over continuous time. Instead, we consider a more difficult problem which consists in inverting for the parameters based only on a finite number of observations. Part of the great richness of inverse problems in general is that the nature of the observations may be constrained in many different ways, often corresponding to practical limitations from applications, each case demanding different solution methods. Here we will focus on discrete observations, and we distinguish two subcases:

– *passive observations* where the glimpse times are not controlled by the observer;
– *active observations* where the observer can choose when the process is glimpsed.

In the latter setting certain constraints still apply; for example, often there is a fixed budget n of available glimpses, or n may be infinite but a fixed average observation rate is imposed.

Our toy example is deterministic. One obtains stochastic scenarios if random measurement errors are considered, or more fundamentally, when replacing the direct equation by a stochastic evolution equation. The inversion problem now becomes one of statistical estimation of the unknown parameters from the observable time series. In the active case, a natural question is that of an optimal spacing of the glimpse times, for example in the sense of minimal estimation variance.

Finally, our toy example and its stochastic versions are *non-intrusive* in that the act of making observations did not perturb the system. A natural extension is to examine the associated intrusive or perturbation problem; for example, each glimpse could add a random impulse to the motion. Would it still be possible to measure the parameters even in this case?

The general aim of this paper is to discuss a class of inverse problems of queueing theory which find their origin in Internet probing.

Over the last 15 years or so a substantial amount of work has been devoted to various types of communication network measurement techniques, all primarily triggered by the need for Internet measurement tools. For instance, the reader can consult [2, 8, 9, 12, 14, 21, 24, 32, 36]. The field of *active probing* of networks, for example, seeks to infer the values of network parameters and the statistics of the teletraffic flowing through it, based on test packets or *probes* which are sent across the network. Here the network model is typically taken to be stochastic, and active (and intrusive) observations are made using a finite number of probes. The *IPPM* Working Group of the IETF and the ACM Sigcomm's Internet Measurement Conference (IMC) provide typical examples of communities where these questions are debated. There are also user associations and communities, for example Grenouille [18] which has hundreds of thousands of members, which rely on end-to-end measurements to determine the service levels, and fairness of access, actually provided by Internet Service Providers (ISPs).

It is our belief that much of what is attempted in these communities can be cast into the framework of inverse problems in queueing theory, or more generally, of inverse problems in discrete event dynamical systems theory. The present paper contains new and recent results in this connection and proposes a classification of questions and problems within this setting. A small number of recent works [1, 5, 6, 26–29, 38] provide rigorous results of this type. The great majority of the literature, however, is focussed on heuristic inversion methods.

Section 2 describes the main concepts of inverse problems in queueing theory and gives a first classification of these problems. The paper is then structured into sections with increasing levels of realism. Section 3 focuses on the case where the observations provide noiseless estimates of certain stationary distributions or moments. This leads to a class of *analytical inverse problems*, where the main output of the method is a closed-form formula or a terminating algorithm providing the exact value of the unknown parameters from the observations.

Section 4 is centered on *statistical inverse problems*, where observations are *finite time series* and where the need is therefore for robust inversion methods taking the noise into account. The main outputs of the method are: (1) a set of estimators that are shown to be asymptotically consistent, and (2) recursive algorithms allowing one to implement the estimation of the unknown parameters from the time series.

Sections 3 and 4 are based on rather specific parametric models which may not be realistic for representing IP networks. The drawback of such parametric methods is that they have to be checked on testbeds and adapted using heuristic modifications in order to cope with real IP networks and traffic (as amply exemplified in e.g. the papers published in the proceedings of the IMC). We will not pursue this line of thought here. We will rather investigate methods which do not suffer from this weakness. This is the object of Sect. 5 which is centered on inversion techniques that work for general classes of models. For these more general systems, we will limit ourselves to the non-intrusive case. In this case, we show that there exist probing strategies leading to asymptotically consistent and minimal variance estimators of the unknown parameters, and this regardless of the specific instance of model taken from this class. The conclusions of Sect. 5 are guidelines and recommendations on how to 'optimally' act in this more general setting. This is linked to the general framework of the *design of experiments* in statistics (see the thesis of B. Parker [31] for the application of this methodology to packet networks).

2 Inverse problems in queueing theory

Our discussion of inverse problems in queueing theory will be from the viewpoint of an Internet prober. That is, an entity whose network observations are derived from probes which are inserted into the network, where the latter is modeled as a queueing system. The default assumption is that only end-to-end measurements on probes are available, that is, that the network does not cooperate in any way and so must be treated as a 'black box'. The reason for this is that Internet service providers are generally either unable, or unwilling, to provide information on their network or the traffic flowing on it. In addition, a route may traverse several Autonomous Systems (administrative domains), implying the need for cooperation across multiple, and competing, providers. Probing is one of the main ways in which knowledge of the growth and performance of the Internet, for example its interconnection graph or topology, is known today. Indeed, service providers themselves use probing despite the fact that they have the option of making measurements directly on their switching infrastructure. The flexible nature of probing, and its direct access to end-to-end metrics important for network applications, makes it an important tool for providers to learn about their own networks. For the end user, it is perhaps their only option. Due to its practical importance, and considerable and growing literature, we focus on this end-to-end probing viewpoint, although of course there exist many other types of inverse problems pertaining to queueing theory. Within the IP network framework there are for instance many interesting *ISP-centric* inverse problems too, which will be briefly discussed in Sect. 2.10. There are also interesting problems in connection with other domains of applications of queueing theory. Let us quote for instance the queue inference engine of R. Larson [25]. This inference engine was designed for ATM machines where the operator of the (cash) machine wants to evaluate the distribution of the customer queue size. The observables are here the epochs of the beginning and the end of all transactions (as recorded by the machine). The busy periods of the single server queue representing the ATM machine can hence be reconstructed from these observations; from this, the law of the queue size can then be evaluated. As we see, the nature of the problem is quite different from what was described above because the observables are quite different (the beginning and end of each service time in the latter case, the arrival times to and the departure times from the queue in the Internet probing case, assuming that one represents the IP path as a single server queue).

2.1 Direct equations of queueing theory

Queueing theory studies the dynamics of stochastic processes in a network of queueing stations, such as queue sizes, losses and delays, as a function of certain parameters. These parameters can be related to the structure of the stations (the number of servers, buffer sizes, service disciplines) or can be the distribution of the stochastic processes driving the queueing network (e.g. the rate of some exogenous Poisson arrival point process, or the law of the service times in a given station). The associated direct equations may bear either on the joint law of these stochastic processes (e.g. the queue sizes form a Markov chain in a Jackson network), or on the recursions satisfied by the random variables themselves (e.g. Lindley's equation for the end-to-end delays for $./GI/1$ FIFO queues in series).

The solution of the direct equation bears on the law of these stochastic processes and might be the steady state or the transient *distribution*. The solution in the recursion viewpoint might be the steady state or the transient random state *random variable*.

In the network probing setting, there are two types of customers in the network: the customers (or packets) sent by regular users, often referred to as *cross-traffic*, and the customers (or probes) sent by the prober performing the measurement experiment. The former are typically fixed, namely the prober has no way to act on the cross-traffic offered to the network, whereas the latter can be sent at will, at least in the case of active probes.

Note that probes are themselves packets. In the active measurement case, their sizes may be chosen at will within a range of values. In the case of the Internet, all IP packets contain a header carrying essential information such as the IP address of the destination, so that 0 size probes are not possible. The maximal size of an IP packet is also fixed, which translates to an upper bound on probe size. In the passive measurement case, probes are just normal packets sent as part of a given application, for instance the packets of a Transport Control Protocol (TCP) flow in charge of a file transfer. The probe sizes are then determined by the selected application and associated network protocol.

A key question within this setting is whether the chosen parametric *queueing model* is an acceptable approximation of the concrete communication network with its cross-traffic and its probes. One most often needs a solution for the direct equation in order to solve the inverse problem. There is hence a crucial trade-off between realism of the queueing model and the mathematical tractability of its direct equation.

2.2 Noise

Deviations from ideal assumptions, which we denote generically by ‘noise’, are present at several levels within this setting:

– Most queueing problems are random by nature: for instance, cross-traffic is best represented as a random process. A key question here is whether the underlying random processes are stationary or not. Since stationarity is most often desirable for tractability, this will lead to upper bounds on the probing period which should not exceed the timescale at which macroscopic, for example diurnal, changes occur.
– There may also be actual measurement noise in the data. In the probing framework, most raw measurements consist of probe departure and arrival timestamps. Neither timestamping, nor the clocks that underlie them, are perfect, and high precision is important in order to resolve small difference in latencies (system times) arising from high capacity links (high service rates). The probability law of the measurement errors can however be well approximated in many cases.
– Finally, there may be noise stemming from the nature of the data itself: all practical time series obtained from a measurement experiment are finite, and so the resulting estimators for parameters are non-degenerate random variables. In others words, there are statistical errors in the parameter estimates.

In spite of all these random phenomena, it may still make sense to consider deterministic direct equations. For instance, the law of a stationary and ergodic stochastic process is a deterministic object, and the pointwise ergodic theorem shows that when the observables contain an infinite time series of samples of such a process, these allow one to reconstruct the stationary law in question with arbitrary precision. In what follows, we will distinguish between noiseless inverse problems, which correspond to a kind of mathematical idealization of reality (e.g. obtained with infinite stationary and ergodic time series, which allow one to determine the exact value of all mean quantities), and noise-aware or robust inverse problems where the intrinsic randomness of the problem is faced.

2.3 Probing actions

The observables are generated through certain actions of the network prober. We below describe what actions are allowed.

Choice of topology Whenever probes traverse more than a single station, the *route* they follow must be specified. Here a route is an input–output/origin–destination pair. Within the IP network setting these end points correspond to interfaces in IP routers. In queueing theory, a natural incarnation is that of a route in the sense of Kelly-type networks [23]. The chief scenarios are as follows. The network probing is:

– *point-to-point* when probes are sent from a single source to a single destination;
– *point-to-multipoint* when probes are sent from a single source to multiple destinations (the network of queues traversed then has a tree-like topology);

– *multipoint-to-point* in the case of multiple sources to a single destination;
– *multipoint-to-multipoint* in the case of multiple sources to multiple destinations.

In the point-to-multipoint case, the actual IP network experiment may differ depending on whether the network has native IP *multicast* available or not. In the former case, probes fork out at each node of the network with a degree larger than 1, and this is well represented by what happens in a Fork–Join queueing network [7]; in the latter case, the experiment will in fact consist of a collection of coordinated point-to-point schemes. In the other cases, the only possibility directly supported by the current Internet is that of a collection of point-to-point schemes. Note that it is generally assumed that all probes traveling from a given source to a given destination pass by the same sequence of internal stations (routers), though this can be generalized.

Passive probing actions Purely passive probing is in fact monitoring, and the only freedom the experimenter enjoys is an ability to filter packets according to various criteria, for example to only take note of TCP packets, and to decide which of these to 'baptize' or *tag* as probes. A less restrictive case is when the prober can in addition control certain overall parameters of probing traffic. In the IP setting, he could for instance select an HTTP application which would initiate several TCP connections whose packets would act as probes, or alternatively a UDP-based application like Voice over IP (VoIP) could be used to generate a probe stream. Here the prober can insure that probes of the desired transport and application type are present, and also decide on when to start and end the flow(s), but there is still no control at the level of individual packet timing.

Active probing actions Active probing consists in sending a set of probes at carefully selected epochs and with carefully chosen sizes. Complete control is possible subject only to constraints on probe size and/or rate as noted above. We include in this category the important case where the probe sizes and their emission times are defined through stochastic processes with fully controlled parameters.

2.4 Observables

Observables are the raw data quantities available to the prober through conducting a probing experiment, and derive from the probing actions just described. In the end-to-end viewpoint, for each route this data consists of probe packet size and departure timestamp at the origin, and loss indication and arriving timestamp (if applicable) at the destination. Effectively therefore, the information is of two types for each route: a loss indication for each probe marking whether it arrived at the destination or not, and if applicable, the probe latency or *delay* in traversing the route.

In the case of active observations, the packet sizes and departure times are in fact controlled by the prober and therefore already known. For simplicity we nonetheless refer to these as observables.

2.5 Unknown parameters and performance metrics

In the context of communication network probing, typical parameters to be identified would be:

– *structure parameters* of the nodes/queueing stations traversed by the probes such as the speed of the link/server, the buffer size, the service discipline used (e.g. to check neutrality, an important requirement of the IETF that packets should not be discriminated against on the basis of the application they stem from);
– *cross-traffic parameters* at a given node if the law of the cross-traffic is in a known parametric class, or otherwise its full distribution.

It is often desirable to estimate certain *performance metrics* such as the packet loss probability, or the distribution of packet latency, along a route or at a given node, in the context of incomplete knowledge of the system parameters.

2.6 Intrusiveness, bias and restitution

Since probes are processed as customers by the queueing system, and moreover have a minimum size which is positive, they interact with cross-traffic and so are inherently intrusive. At first glance, this seems to make the inverse problems more difficult. In fact, as we shall see, intrusiveness may be useful and can be leveraged in many cases (for example, see the poly-phase methods introduced below).

As a result of intrusiveness, in general, the performance metrics of the system with cross-traffic and probes differ from those of the system with cross-traffic only. The performance metrics (or the parameters) of the system "without the probes" are often referred to as the *ground truth* in the network probing literature. For instance, the probability that a typical packet of the cross-traffic on a given route will be lost if there were no probes, or the mean cross-traffic load at the kth router on a route, belong to the ground truth. More generally, the parameters listed above (structural or pertaining to cross-traffic) are by definition part of the ground truth.

An important question is the reconstruction of some ground truth metric from the observation or the estimation of the metric for the perturbed system. This will be referred to as *restitution* below.

Restitution may even be needed in the non-intrusive case (for example, when probes have zero size and system time is the metric of interest) because of the *sampling bias* problem: a typical example is when the ground truth can be evaluated from certain time-averages and where probe-averages do not coincide with time-averages.

2.7 Identifiability, ambiguity

The observables, either implicitly or explicitly, carry information regarding a spatiotemporal slice of the network experienced by the probes. This information is clearly partial, which gives rise to a host of system identifiability questions. For example, in the context of intrusive probing, it is not clear whether the restitution of many ground truth metrics is possible even in principle.

We shall see below that some parameters or performance metrics of a queueing system are not always identifiable from the observables. In some cases, different parameters can lead to the same observations.

2.8 Estimation problems

As mentioned above in Sect. 2.2, in practice the duration of a probing experiment compatible with stationarity is finite, and the number of probes that can be sent during a finite time interval is likewise finite. As a result, in practice the observables consist of time series of finite length, and inversion for the unknown parameters based on them is no longer a deterministic problem, but one of statistical estimation. This leads to a new class of problems in the design of such estimators, and the establishment of their properties, in particular the classical ones of bias, variance, asymptotic consistency and asymptotic normality.

In the case of active probing, the degrees of freedom in how probes are sent allow for another level of problems built on optimizing the statistical properties in the above. For example, a natural question is to ask how probes should be spaced so as to minimize estimation variance.

2.9 The prober's path(s) to ground truth

Let us summarize by stressing that all paths to a given ground truth or performance metric require the following series of steps:

(1) a tractable and yet realistic direct equation for the dynamics of the observables;
(2) a proof of the identifiability of the perturbed metric from the observables;
(3) the definition (and possibly the optimization) of estimators for these metrics;
(4) the design of a restitution mechanism allowing one to reconstruct the ground truth from the perturbed or biased metrics.

The aim of the following sections is to illustrate the above in a few fundamental scenarios. Fortunately enough, some of the requirements may be relaxed in some cases; one may for instance

- idealize step 3, by assuming an infinite time series and therefore, for example, a full knowledge of the stationary distribution of some observable; this leads to deterministic problems that will be illustrated in Sect. 3;
- avoid step 4, by selecting an active probing strategy involving probes rare and small enough to have almost no impact, which justifies a claim that the perturbed and unperturbed systems are the same in practice.

Of course, the validity of such simplifications will have to be discussed in detail.

2.10 ISP-centric inverse queueing problems

The scenarios considered in the remainder of the paper focus on point-to-point inverse problems (which are often more challenging than their multipoint counterparts) arising in active Internet probing with end-to-end observables. For the sake of completeness, we now add a few words on other practical incarnations of inverse problems in queueing theory stemming from the ISP viewpoint.

The simplest observables for an ISP are time series of individual queue sizes and traffic (service times and packet sizes and arrival times) at the input or output ports of its own routers. The ISP has the privileged option of directly and non-intrusively monitoring these, though at the cost of additional monitoring hardware. Its actions then primarily consist in choosing when and what queues or traffic processes to monitor. The parameters and metrics of interest are quite different from, in some sense inverse to, those alluded to in the above. An elegant example is that of the reconstruction of end-to-end metrics, such as the packet-loss point process or the fluctuations of end-to-end delays (jitter) experienced by a typical user whose packets pass by the monitored router, given the node-based observables.

Other aspects of the problem, such as the direct equations to be used, their random nature, the resulting need for estimators of the metrics of interest, are all quite similar to what was described above in the Internet prober case.

3 Noiseless inverse queueing problems

As mentioned above, in this section we assume that the availability of an infinite time series has provided perfect knowledge of the distribution function of the end-to-end stationary observables, so that step 3 from Sect. 2.9 may be skipped. This is an idealization of the noise-aware case, which we study in Sect. 4.

Within this context, we discuss three types of classical models of queueing theory on which Internet probing type inversion is possible: $M/G/1$, $M/M/1$ loss, and Kelly networks. The methods described in this section all leverage the fact that probes are intrusive. They consist in varying the probing rate and in observing how the system reacts to this variation. There are again various levels of realism: one can either assume, as in Sect. 3.1, that the *mapping* that describes the variation of the observation as a function of the probing rate can be deduced from the observations, or pursue a more realistic scenario (considered in the other subsections) where one knows the value of this variation at some finite number of points (probing rates), as in the ‘finite number of glimpses’ scenario of the Introduction.

There is a small literature on this analytic approach, scattered in the communication network literature, particularly the proceedings of venues with a strong Internet focus. Among these the first seems to be [38]. Another early paper advocating an analytical inversion for the estimation of loss processes in networks is [1]. The approach in the latter is moment-based (see below).

3.1 The $M/G/1$ queue

Before probes are injected, the system consists of a FIFO $M/G/1$ queue with a single server with speed 1. The input rate λ and the service distribution G are the two unknown parameters of cross-traffic. The sizes of probes obey a law K (this is the service time for probes) and arrive according to a Poisson point process with rate x. The active prober only has access to the distribution of end-to-end delays of probes. Can he reconstruct the unknown parameters λ and G?

The direct equation is the Pollaczek–Khinchin (PK) formula [40] which stipulates that the stationary waiting times of probes have, for Laplace Transform (LT),

$$\widehat{W}(s) = \frac{(1 - x\overline{K} - \lambda\overline{G})s}{s - x(1 - \widehat{K}(s)) - \lambda(1 - \widehat{G}(s))},$$

where $\widehat{K}(\cdot)$ and $\widehat{G}(\cdot)$ denote the LT of K and G respectively, and $\overline{K}$ and $\overline{G}$ their means. We assume that

$$x\overline{K} + \lambda\overline{G} < 1$$

which is necessary and sufficient for the existence of a stationary regime. Since λ and $\overline{G}$ are unknown, it is impossible to check this condition without prior knowledge. Most Internet resources have a moderate utilization factor (i.e. $\lambda\overline{G}$ rarely exceeds $3/4$ or even $1/2$) and if $x\overline{K} \ll 1$, then the last condition is quite likely to hold. Note that as a general principle probing overhead should be kept small, in order to avoid consuming network bandwidth, to reduce intrusiveness, and to prevent probes being confused with network attacks, so assuming $x\overline{K} \ll 1$ is quite reasonable.

From our infinite-time series assumption, we have access to any function of the stationary end-to-end delay process of probes. In particular, the function $\widehat{W}(s)$ is *indirectly observable* (i.e. can be obtained from the direct delay observable) for all values of x and K since the waiting time of a probe is obtained by subtracting its service time—which is known to the prober—from its end-to-end delay.

We now proceed to *invert* the direct equation. By letting s go to infinity, we have

$$\widehat{W}(\infty) = P_x(W = 0) = 1 - x\overline{K} - \lambda\overline{G} = \kappa(x)$$

which is also indirectly observable. Hence for all x, $1 - \kappa(x) = x\overline{K} + \lambda\overline{G}$. Pick two different values x_1 and x_2 of x. The linear system

$$\begin{aligned} 1 - \kappa(x_1) &= x_1\overline{K} + \lambda\overline{G}, \\ 1 - \kappa(x_2) &= x_2\overline{K} + \lambda\overline{G} \end{aligned} \tag{1}$$

determines λ and $\overline{G}$, which, substituting into the PK transform

$$\widehat{G}(s) = \frac{(1 - x\overline{K} - \lambda\overline{G})s}{\widehat{W}(s)\lambda} - \frac{s - x - \lambda + x\widehat{K}(s)}{\lambda}, \tag{2}$$

determines the transform of the entire law G. Therefore, our two unknown parameters can be unambiguously estimated from such observables.

This approach also allows us to estimate the ground truth stationary end-to-end delay distribution. The restitution formula consists in again applying the PK formula for waiting time, but this time without the probe traffic, which is possible since λ and G are now known.

The main weaknesses of the present approach should be clear:

- it requires the estimation of *distribution functions* (here LTs) rather than *moments*; it may be desirable to have *moment-based* methods (see Sects. 3.2 and 3.4 below);
- it in fact requires several infinite time series, one per value of x; for instance, the solution of the linear system (1) would in practice require two successive *phases*: a phase where the prober sends probes at rate x_1 and collects enough samples to have a precise enough estimate of the stationary probability $P_{x_1}(W = 0)$ that a probe sees an empty system upon arrival; a phase where he sends probes at rate x_2 and obtains estimates for $P_{x_2}(W = 0)$, which is a new system requiring a new time series. It would be desirable to have *mono-phase* inversion techniques.

3.2 The $M/M/1$ queue

The setting of this section is slightly different from that of the last section. The system is a $M/M/1$ FIFO queue with a server of unknown speed μ. Cross-traffic is Poisson with unknown intensity λ and exponential packets with mean 1. The active prober sends Poisson probes with rate x to the system. All probes have exponential size of mean 1. Can one reconstruct λ and μ when observing only the mean stationary end-to-end delays experienced by the probes?

The stationary mean number of packets and probes in the station is

$$\overline{N}(x) = \frac{\lambda + x}{\mu - \lambda - x},$$

under the condition $\lambda + x < \mu$. From Little's formula the mean end-to-end delay D of probes (or packets) is

$$\overline{D}(x) = \frac{\overline{N}}{\lambda + x} = \frac{1}{\mu - \lambda - x}. \tag{3}$$

This formula, which is our direct equation, shows that the constant $\mu - \lambda$, which carries the interpretation of *residual bandwidth*, can be reconstructed from the observation of $\overline{D}$ associated with two different values of x. However, the individual constants μ and λ *cannot* be reconstructed individually from this alone. Fortunately enough, this mean residual bandwidth is sufficient for the restitution of the ground truth cross-traffic delay $\overline{D}(0) = \frac{1}{\mu-\lambda}$.

Let us summarize our conclusions on this case: we have here a first-moment based, though poly-phase (here two-phase) probing strategy allowing one to determine unambiguously the mean residual bandwidth of an $M/M/1$ queue solely from the measurement of the empirical mean end-to-end delays experienced by probes. Within this context, the problem of identifying the intensity of cross-traffic or the speed of the server is however ill-posed.

When adding second-order estimates, one obtains the additional information needed to resolve the two parameters. For instance, when sending packet pairs with size y at the same time, one gets that their system times, D and D', are such that $D' - D = y/\mu$ so that μ can be determined (this packet-pair method actually holds for all $G/G/1$ FIFO queues). In reality, two packets cannot arrive exactly at the same time. It is shown in Appendix 6.1 that in the $M/M/1$ queue, two packets with size y sent t seconds apart have system times D and D' which are such that, as t goes to 0,

$$\mathbb{E}(DD') = K(y) - t\left((1-\rho)y + \frac{\rho}{\mu}\right) + o(t)$$

where $K(y)$ is some constant. The slope w.r.t. t of the function $t \to \mathbb{E}(DD')$ is $(1-\rho)y + \frac{\rho}{\mu}$ and it can be estimated, so that $1-\rho$ and $\frac{\rho}{\mu}$ can also be estimated to arbitrary precision, using different values of y. This determines both λ and μ unambiguously.

There are other practical methods to evaluate μ not based on moments. The simplest one consists in sending probes with constant size y and in looking for the probes with minimal delay. This minimal delay of course allows one to determine μ unambiguously.

3.3 The $M/M/1/B$ queue

The setting is the following: the prober sends Poisson probes with rate x into a system which, without the probes, would be an $M/M/1/B$ queue with Poisson (cross-traffic) input point process of unknown intensity λ. Cross-traffic packets are assumed to have exponential sizes of parameter 1, and the prober emulates this by choosing to send probes with the same size distribution.

Under natural independence assumptions, the full system (with cross-traffic and probes) is an $M/M/1/B$ queue with arrival rate $\lambda + x$ and service rate μ. The direct equation is the following classical expression for the stationary loss probability $p(x)$ (see for example [40]):

$$p(x) = \frac{(\frac{\lambda+x}{\mu})^B - (\frac{\lambda+x}{\mu})^{B+1}}{1 - (\frac{\lambda+x}{\mu})^{B+1}}. \tag{4}$$

Similarly, the probability $q(x)$ that the queue is empty is

$$q(x) = \frac{1 - \frac{\lambda+x}{\mu}}{1 - (\frac{\lambda+x}{\mu})^{B+1}}. \tag{5}$$

Can one determine λ, μ and B, assuming that these parameters (or some of them) are unknown?

From our infinite-time series assumption, we have access to the loss rate $p(x)$ as well as to the sequence of end-to-end delays for each probe. Using packet-pair techniques [32], or alternatively by observing delay minima when probes are chosen of constant size, it is possible to extract the server speed μ. We therefore assume that μ is known. One key consequence of knowing μ is that the prober then knows the service time of each probe, and he can therefore measure the empirical probability $q(x)$ that the queue is empty, since for probes which encounter an empty queue the observed end-to-end delay is equal to the service time.

Assume a poly-phase probing scheme with N different probe intensities x_i, $i = 1, \ldots, N$. Within our noiseless setting, the prober's measurements allow him to determine the associated loss rate p_i and empty queue probability q_i, and hence to compute the ratio $r_i = \frac{p_i}{q_i}$. From (4) and (5), the following should hold for all measured ratios:

$$\forall 1 \le i \le N, \quad r_i = r(x_i) = \frac{p(x_i)}{q(x_i)} = \left(\frac{\lambda + x_i}{\mu}\right)^B,$$

where $r(x)$ is the polynomial $(\lambda + x)^B/\mu^B$. For all $N \ge 1$ let $L_N(x)$ denote the Lagrange polynomial interpolating the points (x_i, r_i), $i = 1, \ldots, N$, namely the polynomial in x of degree at most $N - 1$ defined by the formula

$$L_N(x) = \sum_{i=1}^{N} r_i \prod_{j \ne i} \frac{x - x_j}{x_i - x_j}.$$

For $N \geq B+1$, we have $L_N(x) = r(x)$ for all x. Hence λ and B can be determined as follows:

- B is the degree of $L_N(x)$;
- $(\frac{\lambda}{\mu})^B$ is the constant term of $L_N(x)$ (or $-\lambda$ is the unique real root of $L_N(x)$).

The main limitation of this characterization is that we do not know when $N \geq B+1$, i.e. how many phases are needed. In other words, we have an algorithm which converges to the correct values when letting N go to infinity, but we have no termination criterion for this algorithm. The following lemma and theorem provide such a termination criterion.

Lemma 1 *Consider the set of polynomials $r_{\lambda,\mu,B}(x)$ with B ranging over the positive integers and λ and μ over the positive real line. Two different polynomials of this family intersect in at most 2 points of the positive real line.*

Proof Consider the polynomials $P_1(x) = r_{\lambda_1,\mu_1,B_1}$ and $P_2(x) = r_{\lambda_2,\mu_2,B_2}$. One can assume without loss of generality that $\lambda_1 > \lambda_2$. Let $\delta = \lambda_1 - \lambda_2$. Setting $y = x + \lambda_2$, the equality $P_1(x) = P_2(x)$ now reads

$$(y+\delta)^{B_1} = \frac{\mu_1^{B_1}}{\mu_2^{B_2}} y^{B_2}.$$

If $B_1 \leq B_2$, let $k = B_2 - B_1$. The equality is equivalent to

$$\left(1 + \frac{\delta}{y}\right)^{B_1} y^{-k} = \frac{\mu_1^{B_1}}{\mu_2^{B_2}}.$$

The left-hand term is a decreasing function of y for positive y, and the right-hand term is constant. There is therefore at most 1 solution for positive y and hence for positive x.

If $B_1 > B_2$, let $k = B_1 - B_2$. The equality is equivalent to

$$\left(1 + \frac{\delta}{y}\right)^{B_1} y^{k} = \frac{\mu_1^{B_1}}{\mu_2^{B_2}}.$$

Assume there exist at least 3 positive solutions $0 < y_1 < y_2 < y_3$. Then applying Rolle's theorem to the function $f(y) = (1 + \frac{\delta}{y})^{B_1} y^k$, we get that there are two points $y_4 \in]y_1; y_2[$ and $y_5 \in]y_2; y_3[$ such that $\frac{\partial f(y_4)}{\partial y} = 0$ and $\frac{\partial f(y_5)}{\partial y} = 0$. Now, note that the derivative

$$\frac{\partial f}{\partial y} = y^{k-1}\left(1 + \frac{\delta}{y}\right)^{B_1 - 1}\left(k\left(1 + \frac{\delta}{y}\right) - \frac{B_1 \delta}{y}\right)$$

admits only one zero $y = \frac{\delta B_2}{k}$, which contradicts the existence of 3 solutions $y_1 < y_2 < y_3$. □

Theorem 1 *Assume we have a set of observation points* (x_i, r_i), $i = 1, \ldots, N$, *stemming from an* $M/M/1/B$ *queue with parameters* λ *and* μ. *If* $N > 2$ *and if the Lagrange polynomial* $L_N(x)$ *interpolating the points* (x_i, r_i), $i = 1, \ldots, N$, *can be written as* $(\frac{\widehat{\lambda}+x}{\widehat{\mu}})^{\widehat{B}}$ *for some positive integer* $\widehat{B}$ *and some positive numbers* $\widehat{\lambda}$ *and* $\widehat{\mu}$, *then* $B = \widehat{B}$ *and* $\widehat{\lambda} = \lambda$.

Proof This is a consequence of Lemma 1. The polynomials $L_N(x)$ and $r_{\lambda,\mu,B}(x)$ intersect in $N > 2$ points, and therefore are equal. □

We have hence a termination rule: increase the cardinal N of the set of points (x_i, r_i), $i = 1, \ldots, N$, until the Lagrange polynomial $L_N(x)$ interpolating these points is of the form $(\frac{\widehat{\lambda}+x}{\widehat{\mu}})^{\widehat{B}}$.

We can hence reconstruct the ground truth (on the intensity of cross-traffic and on the loss probability for cross-traffic packets in the absence of probes) by using the formulas for the $M/M/1/B$ queue again, since all the missing parameters are now determined.

More elaborate questions can be addressed along similar lines, for example concerning the determination of the parameters when μ is unknown, but we will not pursue this line of thought here as our aim is more to illustrate the set of problems and solution methods than to provide an exhaustive set of solutions.

3.4 Kelly networks

The systems treated thus far were hardly networks. This subsection is focused on point-to-point probing with many stations, for which we propose a queueing model based on product form theory, and a poly-phase, moment-based inversion method based on a linear interpolation. Since this system will be our reference model in what follows, we provide a discussion, in Appendix 6.2, of how well it maps to real IP networks.

3.4.1 The system

We first describe the system without its probes. It consists of a Kelly network with K stations $S = s_1, \ldots, s_K = D$ and $K + 1$ routes. Route 0 has an exogenous arrival point process which is Poisson of intensity λ_0 and follows the path $s_1, \ldots, s_K$. Route i, for $i \neq 0$ has an exogenous Poisson arrival process of intensity λ_i and its path is the singleton s_i. All packets have exponential size with mean 1. The service rate (or the speed) of s_i is μ_i.

The prober sends probes according to a Poisson point process with rate x and with exponential sizes with mean 1. Probes follow the same path as flow 0 (namely from S to D). We are hence within the context of point-to-point probing.

The unknown parameters are $\lambda_0, \lambda_1, \ldots, \lambda_K, \mu_1, \ldots, \mu_K$. The observables are the stationary end-to-end delays experienced by the probes.

3.4.2 The direct equation

Let us first give the stationary distribution of the end-to-end delays of probes, our direct equation within this setting.

Let us denote by N_i^j and X^j the number of packets of class i, and the number of probes respectively, in station j in steady state. From the product form of Kelly networks [23], we know that if $x + \lambda_0 + \lambda_j < \mu_j$ for all j, then

$$\mathbb{P}\big(X^j = k^j, N_0^j = n_0^j, N_j^j = n_j^j, j = 1, \dots, K\big)$$
$$= \prod_{j=1}^{K} \frac{(n_0^j + n_j^j + k^j)!}{n_0^j! n_j^j! k^j!} \frac{\lambda_0^{n_0^j} \lambda_j^{n_j^j} x^{k^j}}{\mu_j^{n_0^j + n_j^j + k^j}} \frac{\mu_j - \lambda_0 - \lambda_j - x}{\mu_j}. \tag{6}$$

Let $\gamma_j = \mu_j - \lambda_0 - \lambda_j$ denote the residual bandwidth on station j. Direct calculations show that the marginal distribution of the number of probes is

$$\mathbb{P}\big(X^j = k^j, j = 1, \dots, K\big)$$
$$= \sum_{n_1^1 \geq 0} \cdots \sum_{n_K^K \geq 0} \sum_{n_0^1 \geq 0} \cdots \sum_{n_0^K \geq 0} \mathbb{P}\big(X_j = k^j, N_0^j = n_0^j, N_j^j = n_j^j, j = 1, \dots, K\big)$$
$$= \prod_{j=1}^{K} \left(\frac{x}{\gamma_j}\right)^{k^j} \frac{\gamma_j - x}{\gamma_j}. \tag{7}$$

These equations tell us that our system is equivalent, from the point of view of the probes, to a new system with K $M/M/1$ stations in series, without any cross-traffic, and where the server of station j has speed $\gamma_j = \mu_j - \lambda_j - \lambda_0$, namely the residual bandwidth on station j in the initial system. From this point on, we will therefore consider such a network. The fact that residual bandwidths are sufficient to characterize (as well as the best one can hope to determine from) stationary end-to-end delays is in line with what was already observed in the one-station case considered in Sect. 3.2.

The generating function of the total number of probes in the (reduced) system in equilibrium is

$$\psi_N(z) = \prod_{j=1}^{K} \frac{\gamma_j - x}{\gamma_j - xz}. \tag{8}$$

Since probe arrivals are Poisson, PASTA [4] tells us that the distribution of the total number of probes in the system in steady state as given by (7) is the same as that just before a probe arrives. The latter also coincides with the probability distribution of the number of probes in the system just after a probe leaves it (see [4], Chap. 3). In addition, the mean value of $\overline{D}(x)$ of the stationary end-to-end delay of a probe in the network is

$$\overline{D}(x) = \sum_{i=1}^{K} \frac{1}{\gamma_i - x}. \tag{9}$$

The proof of this formula, which is the basis of what follows, can be found in Appendix 6.3.

Remark 1 More general classes of cross-traffic paths can also be considered within this framework. In such an extension, there are as many traffic paths as there are pairs of integers (i, j) with $1 \le i \le j \le K$. A path of type (i, j) brings cross-traffic which is Poisson and enters the network on station i and leaves it from station K. The methodology described above works in this more general setting. It is easy to show that the final result is exactly the same as above, namely (8) and (9) still hold with γ_i now equal to $\mu_i - \xi_i$, where ξ_i denotes the sum of the intensities on all paths traversing node i.

3.4.3 Linear system inversion

In this case, we use a first-moment poly-phase inversion technique, under the following assumption: the prober can measure the mean end-to-end delay of probes for each phase, and the number of stations is known (in real IP networks the latter can be measured by such tools as *traceroute*). We will explain how the prober can compute the coefficients of the polynomial whose roots are the residual bandwidths of each station on the path.

From (9) the mean end-to-end delay can be expressed as follows:

$$\overline{D}(x) = \sum_{i=1}^{K} \frac{1}{\gamma_i - x} = \frac{\sum_{k=0}^{K-1} a_k x^k}{\sum_{k=0}^{K} b_k x^k}, \tag{10}$$

where a_k, b_k are real numbers defined by

$$\sum_{k=0}^{K} b_k x^k = \prod_{i=1}^{K} (\gamma_i - x), \qquad \sum_{k=0}^{K-1} a_k x^k = \sum_{i=1}^{K} \prod_{j \neq i} (\gamma_j - x).$$

So

$$b_k = (-1)^k \sum_{(i_1,\ldots,i_{K-k}), i_j \neq i_l} \gamma_{i_1} \cdots \gamma_{i_{K-k}},$$

$$a_k = (-1)^k (k+1) \sum_{(i_1,\ldots,i_{K-1-k}), i_j \neq i_l} \gamma_{i_1} \cdots \gamma_{i_{K-1-k}} = (-1)(k+1) b_{k+1}.$$

The γ_is are the roots of the denominator polynomial $\sum_{i=0}^{K} b_k x^k$. Therefore, if we identify the b_k variables, we have solved the inverse problem that consists in determining all residual bandwidths from the observations.

We now show how to find the coefficients of the polynomial. Assume we have K perfect measurements $d_j = \overline{D}(x_j)$ of the mean delays for K different values $x_1, \ldots, x_K$ of the probe rate (we will consider the situation with a number of phases larger than K in Sect. 4.1). The method is hence moment-based and poly-phase. We want to find $(b_k)_{k=0,\ldots,K}$ such that

$$\forall j = 1, \ldots, K, \quad d_j = \frac{\sum_{k=0}^{K-1} a_k x_j^k}{\sum_{k=0}^{K} b_k x_j^k} = \frac{\sum_{k=0}^{K-1} -(k+1) b_{k+1} x_j^k}{\sum_{k=0}^{K} b_k x_j^k}. \tag{11}$$

Rational fractions are defined up to a multiplicative factor: we can hence always assume that $b_K = 1$. The system is now equivalent to

$$\forall j = 1, \ldots, K, \quad \sum_{k=0}^{K-1} d_j x_j^k b_k + \sum_{k=1}^{K-1} k x_j^{k-1} b_k = -d_j x_j^K - K x_j^{K-1}, \tag{12}$$

which can be written as the matrix equation $Y = XB$, where X is the $K \times K$ square matrix

$$X_{j,k} = \big((k-1) x_j^{k-2} + d_j x_j^{k-1}\big), \quad j, k = 1, \ldots, K$$

and Y (resp. B) the column vector $Y_j = -K x_j^{K-1} - d_j x_j^K$ (resp. $B_j = b_{j-1}$). When X is invertible, there is only one solution $B = X^{-1} Y$.

We lack sufficient conditions for X to be invertible. The prober will therefore have to continue adding phases until X becomes invertible.

Numerical illustration Table 1 gives some numerical results for this method. The first column indicates the ground truth, i.e. the real values of $(\gamma_1, \ldots, \gamma_K)$. The second column specifies the probing intensities that were used, that is, the vector $(x_1, \ldots, x_K)$. The third column consists of the coefficients of the polynomial $\sum_{i=0}^{K} b_i x^i$, which we write as the vector $B^t = (b_0, \ldots, b_{K-1})$. Finally, the last column gives the estimation of our method, i.e. the values of $(\widehat{\gamma}_1, \ldots, \widehat{\gamma}_K)$. The technique was implemented using Maple, and provides accurate results in all the cases we tried. However, with 7 (or even 5) stations, one can already notice some rounding errors in the calculations. These errors, which stem both from the inversion of the matrix X

Table 1 Linear inversion in Kelly networks: numerical results

Ground truth	Intensities	Vector **B**	Estimation
(10, 30, 70)	(1, 2, 7)	$(-21000, 3100, -110)$	(10, 30, 70)
(10, 25, 30, 60, 70)	(0.3, 1, 2, 4, 7)	$(-3.15 \times 10^7, 6.43 \times 10^6, -4.49 \times 10^5, 1390, -195)$	(10, 25, 29.99, 60.08, 69.92)
(10, 12, 25, 30, 60, 85, 130)	(0.001, 0.3, 1, 2, 4, 7, 9.7)	$(-6 \times 10^{10}, 1.76 \times 10^10, -1.97 \times 10^9, 1.08 \times 10^8, -3.12 \times 10^6, 4.74 \times 10^4, -354)$	(10, 12, 25.05, 29.84, 62.72, 78.78, 135.3)

and the determination of the roots of the polynomial $\sum_{k=0}^{K} b_k x^k$, grow as the number of stations increases.

4 Noise-aware methods

To the best of our knowledge the methods introduced in this section are new in the queueing theory context. They will all be exemplified on the Kelly network model. The section starts with an adaptation of the method of Sect. 3.4.3 to the case of error-prone (i.e., with noise) estimations of end-to-end delays (Sect. 4.1). The latter is moment-based but poly-phase, and fails to find stable estimators. We then present in Sect. 4.2 distribution-based but mono-phase inversion method using *maximum likelihood*. The maximum likelihood method is then revisited in Sect. 4.3 using the *Expectation Maximization* algorithm, which leads to an explicit iteration scheme. Finally, we investigate in Sect. 4.4 the situation where an additive measurement noise has to be taken into account.

4.1 Minimizing quadratic-like error in Kelly networks

The setting is that of Sect. 3.4.3, but we now take into account the fact that the variable d_j in (11) is some error-prone measurement of the stationary mean delays of the probes of phase j. Assuming that the linear system is of full rank, (12) has still one unique solution. However, as shown in Table 2, the method is extremely sensitive to the presence of noise, and solutions are meaningless with as little as 1% error in the measurements.

This sensitivity to noise is due to several reasons: first, the algorithm finds one exact rational fraction, but this fraction interpolates the noised measurements (this is the overfitting phenomenon). Second, the imprecision is multiplied when taking the inverse of X and then when finding the roots of the polynomial. The concatenation of these operations is quite unstable.

In order to prevent the overfitting phenomenon, we explored the classical solution consisting in increasing the number of measurements. Let us assume we have $N > K$ error-prone measures $d_j = \overline{D}(x_j)$ of the mean delays for N different values $x_1, \ldots, x_N$ of the probing rate.

Table 2 Numerical results for linear interpolation. Delays are measured with 1% error (half with 1% more, half with 1% less). Intensities are similar to the ones used in Table 1

Ground truth	Vector **B**	Estimation
(10, 30, 70)	$(6564, -938, 19.9)$	$(-44.4, 10.9, 13.6)$
(10, 25, 30, 60, 70)	$(-14405, 3039, -358, 86, -15.82)$	$(-2.46 - 5.22i, -2.46 + 5.22i, 6.191 - 3.66i, 6.191 + 3.66i, 8.35)$
(10, 12, 25, 30, 60, 85, 130)	$(1.55 \times 10^6, -3.82 \times 10^5, 3100, 1186, -891, 232, -25.5)$	$(-3.42, 0.1 - 4i, 0.1 + 4i, 5.31 - 2.62i, 5.31 + 2.62i, 8.21, 9.91)$

Following the same lines as in Sect. 3.4.3, we arrive at the matrix equation $\tilde{X}B = \tilde{Y}$, where $\tilde{X}$ is the $N \times K$ matrix with (i,k) entry equal to $(k-1)x_i^{k-2} + d_i x_i^{k-1}$, and where $\tilde{Y}$ is the $N \times 1$ vector with i entry $-Kx_i^{K-1} - d_i x_i^K$.

This corresponds to a multiple linear regression, with more measurements than parameters. There is often no unique solution to such a system. A common way to circumvent this difficulty is to select the value $\widehat{B}$ that minimizes the sum of the square errors in each equation:

$$\widehat{B} = \min_B (\tilde{Y} - \tilde{X}B)^t(\tilde{Y} - \tilde{X}B) = \min_{(b_0,\ldots,b_{K-1},1)} \sum_{j=1}^{N} \left[\sum_{k=0}^{K} \left(kx_j^{k-1} + d_j x_j^k \right) b_k \right]^2. \quad (13)$$

The least-squares error solution to (13) is $\widehat{B} = (\tilde{X}^t\tilde{X})^{-1}\tilde{X}^t\tilde{Y}$.

Notice that finding the coefficients b_k which minimize the sum in (13) is not equivalent to minimizing the square of the differences between the left-hand side and the right-hand side of (11). We have in fact multiplied the jth difference by $\sum_{k=0}^{K} b_k x_j^k = \prod_{i=1}^{K}(\gamma_i - x_j)$ before looking for the minimum. The last product is positive and decreasing as x_j increases, so that we put more weight on less intrusive measures. There are several other ways of estimating B (e.g. through total least-square methods [16]) and it would be interesting to compare them. We will not pursue this line of thought as the last step of the inversion method (that consisting in determining the zeros of a polynomial from its coefficients) is in any case likely to be unstable, as illustrated by the following numerical example.

Numerical illustration A Maple implementation indicates that the overfitting correction is not sufficient. We still get complex roots to the polynomial. We conjecture that this is due to the instability when inverting the matrix $\tilde{X}^t\tilde{X}$ and when finding the roots of the polynomial. A small error in the measured delay is amplified by the matrix inversion, and it is well known that a small difference in the coefficients of a polynomial can have a huge impact on its roots. Table 3 provides a few numerical results for the 3-stations case. This instability motivates the maximum likelihood methods studied in the next subsections.

Table 3 Least-squares linear regression in the 3-servers case. The ground truth is (10, 30, 70). Error in mean delay is 1%

N	Intensities	Vector **B**	Estimation
3	(1, 2, 7)	(6563, −938, 19.9)	(−44.4, 10.9, 13.6)
5	(0.3, 1, 2, 4, 7)	(−6075, 914, −39.8)	(9.8, 14.99 − 19.9i, 14.99 + 19.9i)
7	(0.001, 0.3, 1, 2, 4, 7, 9.7)	(−9583, 1417, −55.14)	(9.88, 22.6 − 21.4i, 22.6 + 21.4i)
10	(0.001, 0.3, 0.5, 1, 2, 4, 4.3, 7, 8.7, 9.7)	(−10766, 1610, −62.7)	(9.9, 26.4 − 19.8i, 26.4 + 19.8i)

4.2 Maximum likelihood in Kelly networks

The network and its probes are as in Sect. 3.4. The observables are now a finite time series of probe end-to-end delays and not an exact moment or distribution as in that section. In this section, we will assume that all samples are identically distributed (i.e. we assume stationarity) and *independent*. The latter assumption is of course not true in general as samples collected at two epochs with a finite time difference are in fact (Markov) correlated. However, if inter-probe times are chosen larger in mean than the mixing time of the system, then it is justified to assume independence. The case with dependent samples will be considered in Sect. 5.

The proof of the following lemma, which provides the direct equation to be used what follows, can be found in Appendix 6.3.

Lemma 2 *Let $\phi(d)$ denote the probability density function at $d \geq 0$ of the stationary delay D of a probe in the system. Then*

$$\phi_{\gamma_1,\ldots,\gamma_K}(d) = \left(\prod_{i=1}^{K} \gamma_i'\right) \sum_{i=1}^{K} \frac{e^{-\gamma_i' d}}{\prod_{j\neq i}(\gamma_j' - \gamma_i')}, \tag{14}$$

with $\gamma_i' = \gamma_i - x$.

The problem can hence be viewed as a classical statistical problem, that of fitting distributions of this class.

4.2.1 The one-station case

For $K = 1$, one can somewhat simplify the notation: the speed of the link is μ; the cross-traffic intensity is λ and the probe intensity is x. The system is a FIFO $M/M/1$ queue. The distribution of the delay D of probes is exponential of parameter $\gamma' = \mu - \lambda - x$, namely it admits the density $\phi_\gamma(d) = \gamma' e^{-\gamma' d}$, for all $d \geq 0$. Assume we have several independent delay samples $(d_1, \ldots, d_n)$. Let $\mathbf{d} = (d_1, \ldots, d_n)$. For independent probe delays, the likelihood of the parameter γ is defined as

$$f_{\mathbf{d}}(\gamma) = \prod_{i=1}^{n} \phi_\gamma(d_i) = \gamma'^n e^{-\gamma' \sum_{i=1}^{n} d_i}.$$

The maximum likelihood estimator of the parameter $\widehat{\gamma}$ is the maximum of the likelihood function. This function is positive, and has 0 as a limit when γ' tends to 0 or to ∞. At $\widehat{\gamma}$, we have $\frac{df_{\mathbf{d}}(\gamma)}{d\gamma} = 0$, which is equivalent to

$$n\widehat{\gamma}'^{n-1} e^{-\widehat{\gamma}' \sum_{i=1}^{n} d_i} - \widehat{\gamma}'^n \sum_{i=1}^{n} d_i e^{-\widehat{\gamma}' \sum_{i=1}^{n} d_i} = 0.$$

Hence

$$\widehat{\gamma}' = \frac{n}{\sum_{i=1}^{n} d_i} = \frac{1}{\bar{d}}. \tag{15}$$

The maximum of likelihood for the available bandwidth is hence: $\widehat{\mu}-\widehat{\lambda}=\widehat{\gamma}=\frac{1}{\bar{d}}+x$. This, together with the strong law of large numbers, shows asymptotic consistency: i.e., the estimator converges to the ground truth when the number of probes tends to infinity.

4.2.2 The two-station case

In what follows, we will use the notation γ_i to mean γ_i' for the sake of notational simplification.

We first evaluate the log-likelihood function and then pose the likelihood equation (18). The key results are: (i) the fact that (18) allows one to determine the MLE estimator, and (ii) that the latter is asymptotically efficient (Theorem 2). This convergence is illustrated by simulation results.

The end-to-end delay of a probe is the sum of two independent exponential random variables of parameters γ_1 and γ_2 (see (14)). Its density at $d>0$ is hence

$$\phi_{\gamma_1,\gamma_2}(d)=\gamma_1\gamma_2\frac{e^{-\gamma_1 d}-e^{-\gamma_2 d}}{\gamma_2-\gamma_1}. \tag{16}$$

If $\gamma_2=\gamma_1=\gamma$ (which has essentially no chance of occurring in practice), the density becomes $\gamma^2 d e^{-\gamma d}$, which coincides with the limit $\gamma_2\to\gamma_1$ of (16).

The likelihood function when we have n independent probe delays $(d_1,\dots,d_n)=\mathbf{d}$ is

$$f_{\mathbf{d}}(\gamma_1,\gamma_2)=\prod_{i=1}^{n}\phi_{\gamma_1,\gamma_2}(d_i). \tag{17}$$

We proceed as above by determining the values of the residual capacities that maximize the log-likelihood function $\log f$:

$$\log f_{\mathbf{d}}(\gamma_1,\gamma_2)=n\big(\log(\gamma_1)+\log(\gamma_2)-\log(\gamma_2-\gamma_1)\big)+\sum_{i=1}^{n}\log\big(e^{-\gamma_1 d_i}-e^{-\gamma_2 d_i}\big).$$

At any local extremum, therefore at $(\widehat{\gamma}_1,\widehat{\gamma}_2)$, we have:

$$\begin{aligned}\left.\frac{\partial\log f_{\mathbf{d}}(\gamma_1,\gamma_2)}{\partial\gamma_1}\right|_{\widehat{\gamma}_1,\widehat{\gamma}_2}&=0=\frac{n\widehat{\gamma}_2}{\widehat{\gamma}_1(\widehat{\gamma}_2-\widehat{\gamma}_1)}-\sum_{i=1}^{n}\frac{d_i}{1-e^{-(\widehat{\gamma}_2-\widehat{\gamma}_1)d_i}},\\ \left.\frac{\partial\log f_{\mathbf{d}}(\gamma_1,\gamma_2)}{\partial\gamma_2}\right|_{\widehat{\gamma}_1,\widehat{\gamma}_2}&=0=\frac{-n\widehat{\gamma}_1}{\widehat{\gamma}_2(\widehat{\gamma}_2-\widehat{\gamma}_1)}+\sum_{i=1}^{n}\frac{d_i}{e^{(\widehat{\gamma}_2-\widehat{\gamma}_1)d_i}-1}.\end{aligned} \tag{18}$$

This equation, which is instrumental in determining the MLE numerically, will be referred to as the *likelihood equation* in what follows. Here are important observations: under the natural non-degeneracy assumption satisfied here, the value of $\widehat{\gamma}_1,\widehat{\gamma}_2$

which maximizes the likelihood is a stationary point, namely a solution of the likelihood equation. However, even in this simple two-station case, there may be spurious solutions to this equation, like e.g. local maxima or minima or saddle points. So for locating the global maximum (i.e. the ML estimator) one should first determine all the solutions of the likelihood equation and then determine the solution with maximal likelihood. More can be said on the matter when the number of samples is large. Setting $X = \frac{\widehat{\gamma}_1}{\widehat{\gamma}_2}$ and $Y = (\widehat{\gamma}_2 - \widehat{\gamma}_1)$, (18) now reads:

$$\frac{1}{X} = \frac{Y}{n}\sum_{i=1}^{n}\frac{d_i}{1-e^{-d_i Y}}, \qquad X = \frac{Y}{n}\sum_{i=1}^{n}\frac{d_i}{e^{d_i Y}-1}. \tag{19}$$

Note that $\widehat{\gamma}_2 = \frac{Y}{1-X}$ and $\widehat{\gamma}_1 = \frac{XY}{1-X}$. Multiplying both equations, we get that Y is a solution of the fixed-point equation

$$Y = g(Y) = \left(\left(\frac{1}{n}\sum_{i=1}^{n}\frac{d_i}{1-e^{-d_i Y}}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{d_i}{e^{d_i Y}-1}\right)\right)^{-\frac{1}{2}}. \tag{20}$$

Notice that 0 is always a solution of (20), when extending the right-hand side by continuity. Once a non-zero solution Y of (20) is obtained, X is derived from (19) and this gives a non-degenerate solution to (18). In general, (20) can have either no other solution (than 0), or several other solutions, depending on n and on the sequence of random samples which are chosen. However, the situation simplifies significantly when n is large. Assume that $\gamma_2 > \gamma_1$. Then, by the strong law of large numbers, for all $Y > 0$,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{d_i}{1-e^{-d_i Y}} &= \mathbb{E}\left(\frac{D}{1-e^{-DY}}\right)\\ &= \frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\int_0^\infty \frac{t}{1-e^{-Yt}}\left(e^{-\gamma_1 t}-e^{-\gamma_2 t}\right)dt\\ &= \frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\sum_{k\geq 0}\left(\frac{1}{(\gamma_1+kY)^2}-\frac{1}{(\gamma_2+kY)^2}\right).\end{aligned}$$

Similarly,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{e^{-d_i Y}d_i}{1-e^{-d_i Y}} = \frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\sum_{k\geq 1}\left(\frac{1}{(\gamma_1+kY)^2}-\frac{1}{(\gamma_2+kY)^2}\right).$$

Hence, for n large, (20) is approximately equivalent to

$$\frac{1}{\frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\sqrt{\xi(0)\xi(1)}} - Y = 0 \tag{21}$$

with $\xi(i) = \sum_{k\geq i}(\frac{1}{(\gamma_1+kY)^2} - \frac{1}{(\gamma_2+kY)^2})$. It is easy to show that (21) always admits 0 and $\gamma_2 - \gamma_1$ as solutions. The function on the L.H.S. of (21) is depicted in Fig. 1

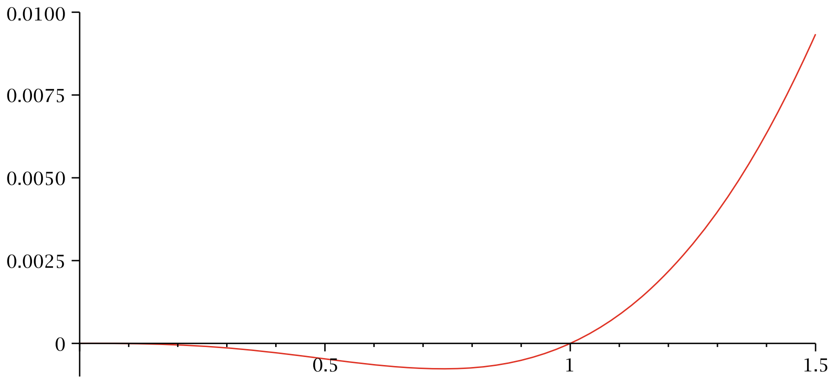

Fig. 1 Shape of the fixed-point equation: L.H.S. of (21)

where one sees that 0 and $\gamma_2 - \gamma_1$ are the only solutions. Hence, we argue that for n large enough, spurious solutions will concentrate around 0 so that $\gamma_2 - \gamma_1$ will be the only other solution.

The main result on this MLE approach is:

Theorem 2 *The MLE* $(\widehat{\gamma}_1, \widehat{\gamma}_2)$ *is asymptotically consistent. That is,* $(\widehat{\gamma}_1, \widehat{\gamma}_2)$ *almost surely converges to the true parameters* (γ_1, γ_2) *when the number of samples n tends to infinity.*

Proof The proof relies on Theorem 7.49 and Lemma 7.54 of [37] which state that if

(1) $\phi_{\psi_1,\psi_2}(d)$ is continuous in (ψ_1, ψ_2) for every d;
(2) $\forall \theta \neq (\gamma_1, \gamma_2), \exists N_\theta$ open set s.t. $\theta \in N_\theta$ and

$$\mathbb{E}_{\gamma_1,\gamma_2}\left[\inf_{\psi \in N_\theta} \log\left(\frac{\phi_{\gamma_1,\gamma_2}(d)}{\phi_{\psi_1,\psi_2}(d)}\right)\right] > -\infty;$$

(3) the parameter space Ω is a compact set,

then the MLE estimator $(\widehat{\gamma}_1, \widehat{\gamma}_2)$ converges almost surely to the true parameters (γ_1, γ_2). In the last expression and below, $\mathbb{E}_{\gamma_1,\gamma_2}[g(d)]$ means integration of the function $g(d)$ w.r.t. the density $\phi_{\gamma_1,\gamma_2}(\cdot)$.

Let us show that our problem verifies the conditions of the theorem. The function $\phi_{\gamma_1,\gamma_2}(d)$ is continuous in (γ_1, γ_2), so that property 1 is verified. By convexity of the exponential function, for all $a < b$ real, $(b-a)xe^{-bx} \le e^{-ax} - e^{-bx} \le (b-a)xe^{-ax}$. Therefore,

$$\gamma_1\gamma_2 d e^{-\gamma_2 d} \le \phi_{\gamma_1,\gamma_2}(d) \le \gamma_1\gamma_2 d e^{-\gamma_1 d}, \tag{22}$$

up to a reordering of γ_1 and γ_2. Therefore, we have:

$$\frac{\gamma_1\gamma_2}{\psi_1\psi_2} e^{(\psi_1-\gamma_2)d} \le \frac{\phi_{\gamma_1,\gamma_2}(d)}{\phi_{\psi_1,\psi_2}(d)}.$$

This implies

$$\inf_{\psi\in N_\theta}\left(\log\left(\frac{\gamma_1\gamma_2}{\psi_1\psi_2}\right)+(\psi_1-\gamma_2)d\right)\le\inf_{\psi\in N_\theta}\log\frac{\phi_{\gamma_1,\gamma_2}(d)}{\phi_{\psi_1,\psi_2}(d)}.$$

Since $\mathbb{E}_{\gamma_1,\gamma_2}[d]=(\frac{1}{\gamma_1}+\frac{1}{\gamma_2})$, we have

$$\begin{aligned}&\log\left(\frac{\gamma_1\gamma_2}{\sup_{\psi\in N_\theta}\psi_1\sup_{\psi\in N_\theta}\psi_2}\right)-\left(\gamma_2-\inf_{\psi\in N_\theta}\psi_1\right)\left(\gamma_1^{-1}+\gamma_2^{-1}\right)\\&\quad\le\mathbb{E}_{\gamma_1,\gamma_2}\left[\inf_{\psi\in N_\theta}\log\left(\frac{\phi_{\gamma_1,\gamma_2}(d)}{\phi_{\psi_1,\psi_2}(d)}\right)\right].\end{aligned}$$

Hence for all bounded open sets N_θ,

$$\mathbb{E}_{\gamma_1,\gamma_2}\left[\inf_{\psi\in N_\theta}\log\left(\frac{\phi_{\gamma_1,\gamma_2}(d)}{\phi_{\psi_1,\psi_2}(d)}\right)\right]>-\infty,$$

so that property 2 is verified. Finally, remember that the parameters are residual bandwidth. Therefore, without losing any meaningful solution, we can restrict the natural parameter space $]0;\infty[^2$ to a space $[\epsilon, A]^2$, where ϵ is a very small capacity (for example, 1 packet per year) and A is the highest capacity of existing routers. □

Remark that Theorem 7.49 of [37] is more general, and that property 3 can be replaced by the following: $\exists C\subseteq\Omega$ compact set s.t. $(\gamma_1,\gamma_2)\in C$ and

$$\mathbb{E}_{\gamma_1,\gamma_2}\left[\inf_{\psi\in\Omega\setminus C}\log\left(\frac{\phi_{\gamma_1,\gamma_2}(d)}{\phi_{\psi_1,\psi_2}(d)}\right)\right]>0,$$

which would allow us to consider any positive value as an acceptable parameter. We are confident that the general form of the theorem holds, and simulations were consistent with this. We choose to use the restricted parameter space because when ϵ and A are well chosen, the restricted parameter space includes all meaningful parameters for the system we consider in practice. Therefore, restricting the parameter space is equivalent to rejecting solutions that we know to be impossible. The question whether the result still holds when taking $\Omega=]0;\infty[^2$ is still open.

We now evaluate the MLE by simulation where delays are generated according to the theoretical law. Residual capacity estimates are obtained using the following technique inspired by the above: we numerically locate the first zero of (20) which is not in the neighborhood of the origin. We use a stopping precision of 10^{-4} in the procedure for finding this zero (a value of 10^{-8} produced the same estimator). In each case, results are averaged over 1000 independent experiments.

Figure 2 plots $\widehat{\gamma_1}$ and $\widehat{\gamma_2}$ when $(\gamma_1,\gamma_2)=(1,2.2)$ as a function of the number of probes n. The results are quite satisfying: for 1000 samples 80% of estimates have error below 10%, and this drops to 4% for 100,000 probes. It is clear that the estimation variance drops, and the right-hand plot shows that it does so as $O(1/n)$, as expected. Notice that $\gamma_2-\gamma_1$ is underestimated. The bias decreases with n also, though this is less obvious in the plots since the decay is much slower than the decay

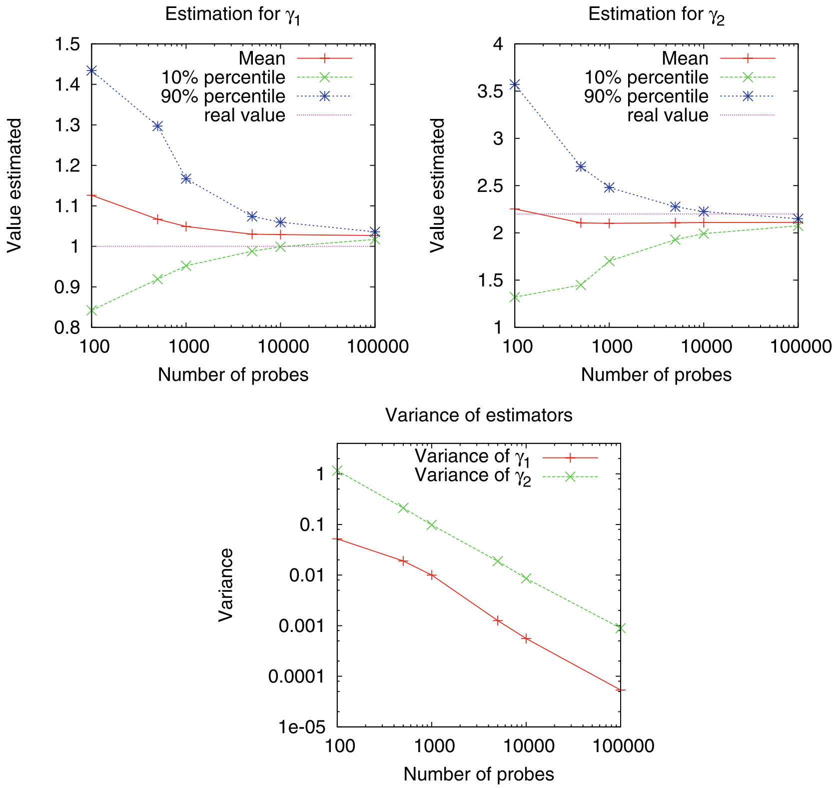

Fig. 2 Precision of the estimated $\gamma_1 = 1$ (*left*) and $\gamma_2 = 2.2$ (*middle*), and variances (*right*, note log scale) as a function of n

of variance. In other words, the MLE is dominated by the bias for large n. If instead we use $(\gamma_1, \gamma_2) = (1, 7.4)$, we approximately obtain the same precision for $\widehat{\gamma_2}$ and improved precision for $\widehat{\gamma_1}$.

It is well known, and to be expected, that the maximum likelihood estimator is biased (although the consistency property implies that asymptotically it is not). For example, in the case of a single server of residual capacity γ and a single probe, the estimator $\widehat{\gamma}$ is simply the inverse of the probe delay D. By convexity of the function $f(x) = \frac{1}{x}$, we get:

$$\mathbb{E}[\widehat{\gamma}] = \mathbb{E}\left[\frac{1}{D}\right] > \frac{1}{\mathbb{E}[D]} = \gamma.$$

4.2.3 More than two stations

This section is focused on the generalization to a path with K routers. We follow the same approach as for the two-stations case. We still use γ_i in place of γ_i'.

According to (14), the likelihood function for n independent end-to-end probe delays $d_1, \dots, d_n$ is

$$f_{\mathbf{d}}(\gamma_1, \dots, \gamma_K) = \prod_{i=1}^{n} \sum_{j=1}^{K} \left(\prod_{k \neq j} \frac{\gamma_k}{\gamma_k - \gamma_j} \right) \gamma_j e^{-\gamma_j d_i}$$
$$= \left(\prod_{p=1}^{K} \gamma_p^n \right) \prod_{i=1}^{n} \sum_{j=1}^{K} \left(\prod_{k \neq j} \frac{1}{\gamma_k - \gamma_j} \right) e^{-\gamma_j d_i}.$$

Therefore, we get the following expression for the log-likelihood:

$$\log\big(f_{\mathbf{d}}(\gamma_1, \dots, \gamma_K)\big) = n \sum_{p=1}^{K} \ln(\gamma_p) + \sum_{i=1}^{n} \log\left(\sum_{j=1}^{K} \left(\prod_{k \neq j} \frac{1}{\gamma_k - \gamma_j} \right) e^{-\gamma_j d_i} \right), \tag{23}$$

so that the likelihood equation reads:

$$\frac{\partial \log(f_{\mathbf{d}}(\gamma_1, \dots, \gamma_K))}{\partial \gamma_l}$$
$$= \frac{n}{\gamma_l} + \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{K} (\prod_{k \neq j} \frac{1}{\gamma_k - \gamma_j}) e^{-\gamma_j d_i}}$$
$$\times \left[\left(e^{-\gamma_l d_i} \prod_{k \neq l} \frac{1}{(\gamma_k - \gamma_l)} \sum_{k \neq l} \frac{1}{(\gamma_k - \gamma_l)} \right) \right.$$
$$\left. - \left(d_i e^{-\gamma_l d_i} \prod_{k \neq l} \frac{1}{(\gamma_k - \gamma_l)} \right) - \left(\sum_{j \neq l} \frac{e^{-\gamma_j d_i}}{\gamma_l - \gamma_j} \prod_{k \neq j} \frac{1}{\gamma_k - \gamma_j} \right) \right].$$

We found no closed-form solution to this system of equation, and instead turn to the Expectation–Maximization algorithm considered below.

4.3 Expectation–maximization algorithm

The Maximum Likelihood estimator is very often analytically difficult or even impossible to derive. One way to overcome this difficulty is to use Expectation–Maximization (E–M). Basically, the E–M algorithm uses additional unknown data in order to find a sufficient statistic and so simplify the maximization problem. The use of E–M algorithm for fitting general phase-type distributions was first described by Asmussen et al. in [3]. The setting considered in the present paper, namely the fitting of sums of independent exponential random variables, is much more specific and this allows us to give explicit iteration formulas and also to prove the convergence of the algorithm, which has not been done for general phase-type distributions to the best of our knowledge.

4.3.1 The two-stations case

In the two-links case, the *incomplete data* are the end-to-end delays d_i of probes, $i = 1, \dots, n$. We *complete* them by the delay on the first link l_i for all probes, $i = 1, \dots, n$, and $\mathbf{l} = (l_1, \dots, l_n)$ denotes their vector. The section starts with the definition of the E–M algorithm in this setting, and then shows that it converges to a solution of the likelihood equation. This proof, which is one of the main mathematical results of the paper, is structured in three lemmas (Lemmas 3 to 5).

The heuristic idea of the E–M algorithm is as follows: The MLE is defined by

$$(\widehat{\gamma}_1, \widehat{\gamma}_2) = \operatorname*{argmax}_{\theta_1, \theta_2} \log\bigl(f_{\mathbf{d}}(\theta_1, \theta_2)\bigr),$$

with $f_{\mathbf{d}}(\theta_1, \theta_2)$ defined in (17), which is not easy to compute. If we knew the complete data $(\mathbf{l}, \mathbf{d})$, it would be quite easy to find

$$\operatorname*{argmax}_{\theta_1, \theta_2} \log\bigl(\widetilde{f}_{\mathbf{l},\mathbf{d}}(\theta_1, \theta_2)\bigr)$$

because this is equivalent to finding

$$\operatorname*{argmax}_{\theta_1, \theta_2} \log\bigl(\widetilde{f}_{\mathbf{l},\widetilde{\mathbf{l}}}(\theta_1, \theta_2)\bigr),$$

where $\widetilde{f}_{\mathbf{l},\mathbf{d}}(\theta_1, \theta_2) = \phi_{\theta_1,\theta_2}(d_1, l_1, \dots, d_n, l_n)$ is the complete likelihood of the complete data and $\widetilde{l}_i = d_i - l_i$ is the delay on the second link, and because the latter is simply

$$\Bigl(\operatorname*{argmax}_{\theta_1} \log\bigl(f_{\mathbf{l}}(\theta_1)\bigr), \operatorname*{argmax}_{\theta_2} \log\bigl(f_{\widetilde{\mathbf{l}}}(\theta_2)\bigr)\Bigr),$$

due to the independence between the vectors $\mathbf{l}$ and $\widetilde{\mathbf{l}}$.

Since the complete data are unknown, one estimates them; more precisely, one starts from some estimate $(\gamma_1^{(k)}, \gamma_2^{(k)})$ of the parameters and, for these parameters, one computes the conditional density

$$\phi_{\gamma_1^{(k)}, \gamma_2^{(k)}}(\mathbf{l}|\mathbf{d})$$

at $\mathbf{l} = (l_1, \dots, l_n)$ of the delays on the first links given the end-to-end delays $(d_1, \dots, d_n) = \mathbf{d}$. One then computes the expectation of the function

$$\mathbf{l} \to \log\bigl(\widetilde{f}_{\mathbf{l},\mathbf{d}}(\theta_1, \theta_2)\bigr)$$

w.r.t. the last conditional density, and then maximizes the last expectation w.r.t. (θ_1, θ_2). This gives a new estimate of the parameters, and one iterates this calculation.

More precisely, let

$$Q_{\mathbf{d}}(\theta_1, \theta_2|\gamma_1, \gamma_2) = \mathbb{E}_{\phi_{\gamma_1,\gamma_2}(\mathbf{l}|\mathbf{d})} \log\bigl(\widetilde{f}_{\mathbf{l},\mathbf{d}}(\theta_1, \theta_2)\bigr). \tag{24}$$

The E–M-algorithm can be defined as follows:

E–M Algorithm: *Take any random* $(\gamma_1^{(0)}, \gamma_2^{(0)})$ *and for each positive integer k, do the following*:

– *Expectation step*: *compute* $Q_{\mathbf{d}}(\theta_1, \theta_2 | \gamma_1^{(k)}, \gamma_2^{(k)})$.
– *Maximization step*: *compute*

$$\big(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}\big) = \underset{(\theta_1,\theta_2)}{\operatorname{argmax}}\, Q_{\mathbf{d}}\big(\theta_1, \theta_2 | \gamma_1^{(k)}, \gamma_2^{(k)}\big). \tag{25}$$

The following lemma illustrates the tractability of this approach:

Lemma 3 *In the two-routers case, for all* $k \geq 0$, (25) *is equivalent to*

$$\frac{1}{\gamma_1^{(k+1)}} = \frac{1}{n}\sum_{i=1}^{n} \frac{d_i e^{(\gamma_2^{(k)} - \gamma_1^{(k)})d_i}}{e^{(\gamma_2^{(k)} - \gamma_1^{(k)})d_i} - 1} - \frac{1}{\gamma_2^{(k)} - \gamma_1^{(k)}}, \tag{26}$$

and

$$\frac{1}{\gamma_2^{(k+1)}} = \frac{1}{\gamma_2^{(k)} - \gamma_1^{(k)}} - \frac{1}{n}\sum_{i=1}^{n} \frac{d_i}{e^{(\gamma_2^{(k)} - \gamma_1^{(k)})d_i} - 1}. \tag{27}$$

Proof We have

$$\begin{aligned}\phi_{\gamma_1,\gamma_2}(l|d) &= \frac{\phi_{\gamma_1,\gamma_2}(l,d)}{\phi_{\gamma_1,\gamma_2}(d)} = \frac{\gamma_1\gamma_2 e^{-\gamma_1 l} e^{-\gamma_2(d-l)}}{\gamma_1\gamma_2 \frac{e^{-\gamma_1 d} - e^{-\gamma_2 d}}{\gamma_2 - \gamma_1}} \\ &= \frac{(\gamma_2 - \gamma_1) e^{(\gamma_2-\gamma_1)l}}{e^{(\gamma_2-\gamma_1)d} - 1},\end{aligned} \tag{28}$$

so that

$$\phi_{\gamma_1,\gamma_2}(\mathbf{l}|\mathbf{d}) = \frac{(\gamma_2 - \gamma_1)^n e^{(\gamma_2-\gamma_1)\sum_{i=1}^n l_i}}{\prod_{i=1}^n (e^{(\gamma_2-\gamma_1)d_i} - 1)}. \tag{29}$$

The expectation step gives:

$$\begin{aligned}&Q_{\mathbf{d}}(\theta_1, \theta_2 | \gamma_1, \gamma_2) \\ &\quad = \sum_{i=1}^{n} \int_0^{d_i} \log\big(\theta_1\theta_2 e^{-\theta_2 d_i} e^{(\theta_2-\theta_1)l_i}\big) \frac{(\gamma_2 - \gamma_1)e^{(\gamma_2-\gamma_1)l_i}}{e^{(\gamma_2-\gamma_1)d_i} - 1} dl_i \\ &\quad = \sum_{i=1}^{n} \log(\theta_1) + \log(\theta_2) - \theta_2 d_i - \frac{\theta_2 - \theta_1}{\gamma_2 - \gamma_1} + \frac{(\theta_2 - \theta_1)d_i e^{(\gamma_2-\gamma_1)d_i}}{e^{(\gamma_2-\gamma_1)d_i} - 1},\end{aligned} \tag{30}$$

so that

$$\frac{\partial Q_{\mathbf{d}}(\theta_1, \theta_2|\gamma_1, \gamma_2)}{\partial \theta_1} = \frac{n}{\theta_1} + \frac{n}{\gamma_2 - \gamma_1} - \sum_{i=1}^{n} \frac{d_i e^{(\gamma_2-\gamma_1)d_i}}{e^{(\gamma_2-\gamma_1)d_i} - 1},$$

and

$$\frac{\partial Q_{\mathbf{d}}(\theta_1,\theta_2|\gamma_1,\gamma_2)}{\partial\theta_2}=\frac{n}{\theta_2}-\frac{n}{\gamma_2-\gamma_1}+\sum_{i=1}^{n}\frac{d_i}{e^{(\gamma_2-\gamma_1)d_i}-1}.$$

The announced result then follows from the maximization step. □

Two important remarks are in order:

– For all k, $\frac{1}{\gamma_1^{(k+1)}}>0$ and $\frac{1}{\gamma_2^{(k+1)}}>0$. This follows from the fact that

$$e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}-1<\big(\gamma_2^{(k)}-\gamma_1^{(k)}\big)d_i e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}.$$

Therefore,

$$\frac{d_i e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}}{e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}-1}>\frac{1}{\gamma_2^{(k)}-\gamma_1^{(k)}}$$

and (26) show that $\frac{1}{\gamma_1^{(k+1)}}>0$. Similarly,

$$e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}-1>\big(\gamma_2^{(k)}-\gamma_1^{(k)}\big)d_i.$$

Therefore,

$$\frac{d_i}{e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}-1}<\frac{1}{\gamma_2^{(k)}-\gamma_1^{(k)}}$$

and (27) imply $\frac{1}{\gamma_2^{(k+1)}}>0$.

– For all $k\ge 0$,

$$\frac{1}{\gamma_1^{(k+1)}}+\frac{1}{\gamma_2^{(k+1)}}=\frac{1}{n}\sum_{i=1}^{n}d_i. \tag{31}$$

This is immediate when adding up (26) and (27).

Here are now the main results on the E–M algorithm in this case.

Lemma 4 *The sequence* $\log f_{\mathbf{d}}(\gamma_1^{(k)},\gamma_2^{(k)})$ *is increasing and converges to a finite limit.*

An obvious corollary of this lemma is that the sequence $f_{\mathbf{d}}(\gamma_1^{(k)},\gamma_2^{(k)})$ is also increasing and converges to a finite limit. The proof, which is based on a classical information theoretic inequality, can be found in e.g. [37]; we recall the main steps in Appendix 6.4.1 as some parts of this proof are needed in what follows.

The fact that the sequence $f_{\mathbf{d}}(\gamma_1^{(k)},\gamma_2^{(k)})$ converges does not prove yet that $(\gamma_1^{(k)},\gamma_2^{(k)})$ converges, and even if it does so, it could converge to some value which is not a solution of the likelihood equation. However, for this particular case:

Lemma 5 *The sequence* $(\gamma_1^{(k)}, \gamma_2^{(k)})$ *converges to a finite limit* (γ_1^*, γ_2^*) *which is a solution of the likelihood equation.*

Proofs for each of these statements appear in Appendix 6.4. Note that we provide an original proof based on a continuity argument, because the natural sufficient conditions adapted from [13] for the convergence of $(\gamma_1^{(k)}, \gamma_2^{(k)})$ do not hold here.

As a direct corollary of these lemmas, if the likelihood equation has a unique solution which is a maximum, then this is the maximal likelihood estimator and the E–M algorithm converges to it, which itself converges to the ground truth as n increases (Theorem 2).

4.3.2 More than two stations

Denote by $l_{j,i}$ the time spent by probe i on link j. If there is only one probe, we just write l_j for the time it spends on link j. Hence

$$\begin{aligned}
\phi_{\gamma_1,\dots,\gamma_K}(l_1,\dots,l_{K-1}|d) &= \frac{\phi_{\gamma_1,\dots,\gamma_K}(l_1,\dots,l_{K-1},d)}{\phi_{\gamma_1,\dots,\gamma_K}(d)} \\
&= \frac{\gamma_1\cdots\gamma_K e^{-\gamma_1 l_1}\dots e^{-\gamma_{K-1}l_{K-1}}e^{-\gamma_K(d-l_1-\dots-l_{K-1})}}{\gamma_1\cdots\gamma_K\sum_{j=1}^K(\prod_{k\neq j}\frac{1}{\gamma_k-\gamma_j})e^{-\gamma_j d}} \\
&= \frac{e^{-\gamma_K d}\prod_{j=1}^{K-1}e^{(\gamma_K-\gamma_j)l_j}}{\sum_{j=1}^K(\prod_{k\neq j}\frac{1}{\gamma_k-\gamma_j})e^{-\gamma_j d}} \\
&= \frac{\prod_{j=1}^{K-1}e^{(\gamma_K-\gamma_j)l_j}}{\sum_{j=1}^K(\prod_{k\neq j}\frac{1}{\gamma_k-\gamma_j})e^{(\gamma_K-\gamma_j)d}}. \qquad (32)
\end{aligned}$$

Then, for a sample of n independent probe delays, we have (with the same notation as above):

$$Q_{\mathbf{d}}(\theta_1,\dots,\theta_K|\gamma_1,\dots,\gamma_K) = \sum_{i=1}^n \mathbb{E}\big[\log\big(\widetilde{f}_{(l_{(1,i)},\dots,l_{(K-1,i)},d_i)}(\theta_1,\dots,\theta_K)\big)\big]$$

where the expectation bears on the variables $l_{(1,i)},\dots,l_{(K-1,i)}$ and is with respect to the conditional density

$$\phi_{\gamma_1,\dots,\gamma_K}(l_{(1,i)},\dots,l_{(K-1,i)}|d_i).$$

This leads to the following integral expression:

$$\begin{aligned}
&Q_{\mathbf{d}}(\theta_1,\dots,\theta_K|\gamma_1,\dots,\gamma_K) \\
&= \sum_{i=1}^n \int_{l_{(1,i)}=0}^{d_i}\cdots\int_{l_{(K-1,i)}=0}^{d_i-\sum_{j=1}^{K-2}l_{(j,i)}} \frac{\prod_{j=1}^{K-1}e^{(\gamma_K-\gamma_j)l_{(j,i)}}}{\sum_{j=1}^K(\prod_{k\neq j}\frac{1}{\gamma_k-\gamma_j})e^{(\gamma_K-\gamma_j)d_i}}
\end{aligned}$$

$$\times \log\left(\theta_1 \cdots \theta_K e^{-\theta_K d_i} \prod_{j=1}^{K-1} e^{(\theta_K-\theta_j)l_{(j,i)}}\right) dl_{(1,i)} \cdots dl_{(K-1,i)}$$

$$= \sum_{i=1}^{n} \alpha_i(d_i) \int_{l_{(1,i)=0}}^{d_i} \cdots \int_{l_{(K-1,i)=0}}^{d_i - \sum_{j=1}^{K-2} l_{(j,i)}} \prod_{j=1}^{K-1} e^{(\gamma_K-\gamma_j)l_{(j,i)}}$$

$$\times \left(\log(\theta_K) - \theta_K d_i + \sum_{j=1}^{K-1} \big(\log(\theta_j) + (\theta_K - \theta_j) l_{(j,i)}\big)\right) dl_{(1,i)} \cdots dl_{(K-1,i)},$$

with

$$\alpha_i(d_i) = \frac{1}{\sum_{j=1}^{K} (\prod_{k\neq j} \frac{1}{\gamma_k - \gamma_j}) e^{(\gamma_K-\gamma_j)d_i}}. \tag{33}$$

These integrals show that $Q_{\mathbf{d}}(\theta_1,\ldots,\theta_K|\gamma_1,\ldots,\gamma_K)$ is an affine function of the variables θ_j and $\log\theta_j$ $\forall 1 \le j \le K$. This means that taking its partial derivative with respect to any θ_j and setting it to zero will give a simple equation of the form $\frac{a}{\theta_j} + b = 0$ to solve, which will provide the solution of the maximization step in closed form. Let us illustrate this by:

Lemma 6 *For the three-routers case, for all $k \ge 0$, (25) is equivalent to*

$$\frac{1}{\gamma_1^{(k+1)}} = -\frac{1}{\gamma_2^{(k)} - \gamma_1^{(k)}} + \frac{\gamma_3^{(k)} - \gamma_2^{(k)}}{\gamma_3^{(k)} - \gamma_1^{(k)}}$$

$$\times \frac{1}{n}\sum_{i=1}^{n} \frac{(\gamma_3^{(k)} - \gamma_1^{(k)}) d_i e^{-\gamma_1^{(k)} d_i} + e^{-\gamma_3^{(k)} d_i} - e^{-\gamma_1^{(k)} d_i}}{(\gamma_3^{(k)} - \gamma_2^{(k)}) e^{-\gamma_1^{(k)} d_i} - (\gamma_3^{(k)} - \gamma_1^{(k)}) e^{-\gamma_2^{(k)} d_i} + (\gamma_2^{(k)} - \gamma_1^{(k)}) e^{-\gamma_3^{(k)} d_i}}, \tag{34}$$

$$\frac{1}{\gamma_2^{(k+1)}} = \frac{1}{\gamma_2^{(k)} - \gamma_1^{(k)}} - \frac{\gamma_3^{(k)} - \gamma_1^{(k)}}{\gamma_3^{(k)} - \gamma_2^{(k)}}$$

$$\times \frac{1}{n}\sum_{i=1}^{n} \frac{(\gamma_3^{(k)} - \gamma_2^{(k)}) d_i e^{-\gamma_2^{(k)} d_i} + e^{-\gamma_3^{(k)} d_i} - e^{-\gamma_2^{(k)} d_i}}{(\gamma_3^{(k)} - \gamma_2^{(k)}) e^{-\gamma_1^{(k)} d_i} - (\gamma_3^{(k)} - \gamma_1^{(k)}) e^{-\gamma_2^{(k)} d_i} + (\gamma_2^{(k)} - \gamma_1^{(k)}) e^{-\gamma_3^{(k)} d_i}} \tag{35}$$

and

$$\frac{1}{\gamma_3^{(k+1)}} = \frac{1}{n}\sum_{i=1}^{n} d_i - \frac{1}{\gamma_2^{(k+1)}} - \frac{1}{\gamma_1^{(k+1)}}. \tag{36}$$

Proof The proof can be found in Appendix 6.5. □

Table 4 provides simulation results for the 3-stations case.

4.4 Additive measurement noise

We consider now the case with additive noise in measurements. We return to the single-station case but we now assume that all delays have some measurement noise

Table 4 Precision of estimator $(\widehat{\gamma}_1, \widehat{\gamma}_2, \widehat{\gamma}_3)$ for various ground truths. Experiments have 10^4 probes and are repeated 200 times

Ground truth	Mean	10% percentile	90% percentile	Variance
(1, 10, 100)	(1, 9.99, 101)	(0.98, 9.11, 78.9)	(1.02, 10.9, 129)	$(1.5 \times 10^{-4}, 0.43, 320)$
(1, 10, 20)	(1, 9.83, 22.2)	(0.99, 7.93, 15.7)	(1.02, 11.9, 30.6)	$(1.5 \times 10^{-4}, 2.3, 35)$
(1, 10, 11)	(1, 8.35, 14.4)	(0.99, 6.83, 11.02)	(1.02, 9.87, 18.5)	$(1.5 \times 10^{-4}, 1.35, 9.43)$
(1, 100, 110)	(1, 68.7, 188)	(0.99, 59.4, 165)	(1.01, 77.7, 213)	$(1.1 \times 10^{-4}, 52.5, 418)$
(1, 2, 100)	(1, 2.01, 91.4)	(0.97, 1.88, 72.1)	(1.04, 2.15, 111)	$(7 \times 10^{-4}, 0.01, 223)$
(1, 1.2, 100)	(1, 1.2, 89,7)	(0.93, 1.1, 72.2)	(1.08, 1.32, 107)	$(3.3 \times 10^{-3}, 6.8 \times 10^{-3}, 212)$
(1, 1.2, 10)	(1.07, 1.09, 13)	(1.05, 1.07, 9.85)	(1.08, 1.1, 17.1)	$(1.4 \times 10^{-4}, 1.7 \times 10^{-4}, 8.19)$
(1, 1.2, 1.4)	(1.04, 1.105, 1.48)	(1, 1.04, 1.36)	(1.1, 1.2, 1.67)	$(1.2 \times 10^{-3}, 3 \times 10^{-3}, 0.015)$

which consists in adding an independent random variable which is uniform in $[0, a]$. The density of the noised delay D is then

$$\phi_\gamma(d) = \begin{cases} \int_0^d \frac{1}{a}\gamma' e^{-\gamma'(d-x)}\, dx = \frac{1-e^{-\gamma' d}}{a} & \text{if } 0 \le d < a, \\ \int_0^a \frac{1}{a}\gamma' e^{-\gamma'(d-x)}\, dx = \frac{e^{-\gamma' d}(e^{\gamma' a}-1)}{a} & \text{if } d \ge a. \end{cases}$$

The likelihood to measure n delays $d_1 \le d_2 \le \cdots \le d_{k-1} < a \le d_k \le \cdots \le d_n$ is:

$$f_{\mathbf{d}}(\gamma) = \frac{1}{a^n}\prod_{i=1}^{k-1}\left(1 - e^{-\gamma' d_i}\right) e^{-\gamma' \sum_{i=k}^{n} d_i}\left(e^{\gamma' a} - 1\right)^{n-k+1}.$$

Direct calculations give that

$$\frac{\partial \log f_{\mathbf{d}}(\gamma)}{\partial \gamma} = -\sum_{i=1}^{n} d_i + \sum_{i=1}^{k-1}\frac{d_i}{1 - e^{-\gamma' d_i}} + (n-k+1)\frac{a}{1 - e^{-\gamma' a}}.$$

Hence, the maximum likelihood estimator $\widehat{\gamma}$, which verifies the relation

$$\frac{\partial \log f_{\mathbf{d}}(\gamma)}{\partial \gamma}(\widehat{\gamma}) = 0,$$

is such that

$$\sum_{i=1}^{k-1}\frac{d_i}{1 - e^{-\widehat{\gamma}' d_i}} + (n-k+1)\frac{a}{1 - e^{-\widehat{\gamma}' a}} = \sum_{i=1}^{n} d_i. \tag{37}$$

The function

$$\widehat{\gamma}' \to \sum_{i=1}^{k-1}\frac{d_i}{1 - e^{-\widehat{\gamma}' d_i}} + (n-k+1)\frac{a}{1 - e^{-\widehat{\gamma}' a}}$$

is increasing. There is therefore only one solution to (37), which can easily be found using numerical techniques. It is easy to check that

$$\begin{aligned}
&\mathbb{E}_{\gamma'}\left[\frac{D}{1-e^{-\gamma' D}}1_{D<a}\right]+\frac{a}{1-e^{-\gamma' a}}\mathbb{P}_{\gamma'}(D\geq a)\\
&\quad=\int_0^a \frac{t}{1-e^{-\gamma' t}}\frac{1-e^{-\gamma' t}}{a}\,dt+\frac{a}{1-e^{-\gamma' a}}\int_a^\infty \frac{e^{-\gamma' t}(e^{\gamma' a}-1)}{a}\,dt\\
&\quad=\int_0^a \frac{t}{a}\,dt+e^{-\gamma' a}\int_a^\infty e^{-\gamma' t}\,dt=\frac{a}{2}+\frac{1}{\gamma'}=\mathbb{E}_{\gamma'}[D]. \qquad (38)
\end{aligned}$$

Hence (38) is equivalent to (37) when the number of probes n tends to infinity. This shows the asymptotic consistency of MLE estimator for one station and uniform noise.

In practice, timestamps are measured at the departure and the arrival of packets. Assuming that timestamps suffer from a uniformly distributed noise, the measured delay is the real delay plus two independent uniform noise variables. The design of maximum likelihood techniques for such noise structures and working for several station in series is an interesting open question.

5 Optimal probing strategies

We have already pointed out that in the error-prone case, once statistical estimators of parameters have been derived based on a given probing stream, one could consider going further by asking how their performance can be optimized by taking advantage of the free parameters of active probing. The difficulty here is that exploring richer probing streams, for example moving away from Poisson probing, implies dealing with more complex direct equations.

In this section we show how taking a more general point of view can lead to insight into the nature of probing streams which are likely to lead to good properties for the associated estimators, such as low estimation variance. To simplify the problem, we focus on the case of non-intrusive probes which have no impact on the system, namely the network and its cross-traffic.

Section 5.1, which builds upon ideas discussed in [5], bears on a question which is often referred to as the *sampling bias* problem and which in fact addresses the issue of the asymptotic consistency of empirical mean estimators.

Section 5.2 bears on the minimization of variance within this context. The main ideas stem from [6].

Section 5.3 discusses a few open problems in the case of maximum likelihood estimators.

5.1 Sampling bias

Consider the following non-intrusive variant of the problem considered in Sect. 4.2.1. The network consists of a single station with cross-traffic consisting in a Poisson point process (with intensity λ) of exponentially sized packets (with mean service time μ). One wants to estimate the residual bandwidth $\mu - \lambda$.

For this, one sends probes of zero size to this system according to some stationary point process which is not necessary Poisson. Let $N = \{T_n\}_n$ denote the points of this point process and let $\{W(t)\}_t$ denote the stationary workload process in the station (since probes have 0 size, this workload is also the ground-truth workload). We will assume this stochastic process to be right-continuous. For all n, let $D_n = W(T_n)$. Since the system is FIFO and all probes have 0 size, D_n is the end-to-end delay measured from probe n. If N and $\{W(t)\}_t$ are jointly stationary, then the sequence $\{D_n\}_n$ is stationary too. If in addition N and $\{W(t)\}_t$ are jointly ergodic, then the pointwise ergodic theorem implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} D_n = \mathbb{E}_N^0\big[W(0)\big], \quad \text{a.s.} \tag{39}$$

In the last equation, E_N^0 denotes expectation w.r.t. the Palm probability P_N^0 of the point process N (see [4]). But if N and $\{W(t)\}_t$ are independent, then $E_N^0[W(0)] = E[W(0)]$, namely probe averages see time averages. Hence, under our assumptions,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} D_n = \frac{1}{\mu - \lambda} \tag{40}$$

a.s., so that we then always have an asymptotically consistent estimator for the residual bandwidth.

Assume now that the network and its cross-traffic form a $G/G/1$ queue with a server with speed 1 and packets with size distributed according to some probability law F on the positive real line. Let $\{W(t)\}$ denote the workload process in this queue. Assume one sends non-intrusive probes according to the point process N. If we have joint stationarity and ergodicity of the two last processes, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1(D_n = 0) = P_N^0\big[W(0) = 0\big], \tag{41}$$

a.s. If N and $\{W(t)\}$ are independent, then $P_N^0[W(0) = 0] = P[W(0) = 0]$. But for all $G/G/1$ queues, $P[W(0) = 0] = 1 - \rho$, where ρ is the load factor of the queue. Hence, under the foregoing assumptions, we have an asymptotically consistent estimator for the load factor, which holds for all $G/G/1$ systems.

Until relatively recently, whenever the ground truth was some time average (or some function of a time average as in the above where the available bandwidth is the inverse of the mean stationary workload), it was recommended to use Poisson probes, namely probes sent at the epochs of a Poisson point process [33, 39]. The rationale for that was that since *Poisson Arrivals See Time Averages* [4], the samples of the metrics estimated by Poisson probes allow one to estimate this ground truth.

The arguments used above show that there is in fact no fundamental reason for using Poisson probes in the non-intrusive case and that a wide variety of other probing strategies share the same ‘lack of sampling bias’, or more precisely asymptotic consistency property.

Let us list and discuss the key assumptions of the last derivation so as to reach a general statement. We consider some system with a continuous time state $\{W(t)\}_t$ assumed to be stationary and ergodic and where the unknown parameters can be determined from the knowledge of $E[W(0)]$. If the prober chooses some probing-point process $N = \{T_n\}$ which is

(1) non-intrusive;
(2) stationary;
(3) independent of $\{W(t)\}$;
(4) jointly ergodic with $\{W(t)\}$,

and if he can observe the quantities $D_n = W(T_n)$, then the empirical mean of the observations is an asymptotically consistent estimator of $E[W(0)]$ and hence of the unknown parameters.

All the above assumptions are necessary. For instance, in the $G/G/1$ queue example, assumption 3 does not hold when N is the point process of all or some selected arrivals of the cross-traffic. In this case (which could be seen as an embodiment of passive measurement), the empirical mean converges but to $E_N^0[W(0)]$ which is then different from $E[W(0)]$ in general. As for assumption 4, if for instance N and $\{W(t)\}$ are both periodic, then there is no joint ergodicity (we have a phase lock) and empirical averages converge to random variables that depend on some random phase. In none of theses cases do we have an asymptotically consistent estimator of $E[W(0)]$.

It is easy for the prober to build a stationary point process independent of $\{W(t)\}$, for instance by making use of a stationary renewal process. A simple way to guarantee assumption 4 is to require that this point process be mixing. Indeed, the product of a mixing and an ergodic shift is ergodic [35].

Hence the general *NIMASTA recommendation*: Non-Intrusive and Mixing probing Arrivals See Time Averages. Poisson processes are mixing and there is no harm using such processes within this setting. But the class of 'good' probing-point processes is much larger as we see.

The property that the sampling of an ergodic stochastic process at the epochs of a mixing and independent point process leads to no sampling bias was first proved in [15].

We conclude this section with a few observations:

– Consider the above framework. If $\{W(t)\}$ is known to be mixing, then all stationary ergodic point processes which are independent of $\{W(t)\}$ lead to an empirical mean estimator of the mean value $E[W(0)]$ which is asymptotic consistent.
– In the intrusive case and when the inversion method is based on the empirical mean estimator of the mean value $E[D_x(0)]$ of some characteristic of the system with its cross-traffic and its probes, Poisson probing is a natural choice as it guarantees asymptotic consistency, as a consequence of the PASTA property.

5.2 Variance

The setting is the same as that of the last subsection, with N a stationary point process with intensity μ. We denote the mean value to be estimated by $p = E[W(0)]$ and we denote the auto-covariance function of $\{W(t)\}_t$ by

$$R(\tau) = \mathbb{E}\big[W(t)W(t+\tau)\big] - p^2.$$

We assume that the function $\tau \to R(\tau)$ exists and is *convex* for $\tau \geq 0$.

The sample mean estimator of p using K samples is

$$\hat{p}_1 = \frac{1}{K}\sum_i W(T_i). \tag{42}$$

The underlying probability is the Palm probability of N. So $T_0 = 0$ by convention and T_i is the sum of i inter-sample times which, due to stationarity, each have law F with mean μ^{-1}. Hence T_i has mean $i\mu^{-1}$, and we denote its law by f_i.

Using the independence assumptions, we get that the variance of $\hat{p}_1$ (which coincides with its mean square error as the estimator is unbiased) is given by

$$\begin{aligned}\mathrm{Var}[\hat{p}_1] &= \frac{1}{K^2}\bigg(K\mathbb{E}\big[W(0)^2\big] + 2\sum_{i\neq j}\mathbb{E}\big[W(T_i)W(T_j)\big]\bigg) - p^2 \\ &= \frac{1}{K^2}\bigg(K\mathbb{E}\big[W(0)^2\big] + 2\sum_{i\neq j}\int R(\tau)f_{|i-j|}(d\tau)\bigg) + p^2\bigg(1 - \frac{2}{K}\bigg).\end{aligned} \tag{43}$$

As a special case of (42), we pick out the estimator based on periodic samples of period μ^{-1}, namely

$$\hat{p}_2 = \frac{1}{K}\sum_i W\big(i\mu^{-1}\big), \tag{44}$$

for which the integral $\int R(\tau)f_{|i-j|}(d\tau)$ in (43) degenerates to $R(|i-j|\mu^{-1})$.

Theorem 3 *Under the above convexity assumption*, $\mathrm{Var}[\hat{p}_1] \geq \mathrm{Var}[\hat{p}_2]$.

Proof Equation (43) holds for all processes. So, to compare the variances it is enough to compare, for all $i \neq j$, the cross-terms, namely $\int R(\tau)f_{|i-j|}(d\tau)$ and $R(|i-j|\mu^{-1})$. But, if $R(\tau)$ is convex, Jensen's inequality says that

$$\int R(\tau)f_k(d\tau) \geq R\bigg(\int \tau f_k(d\tau)\bigg) = R\big(k\mu^{-1}\big), \tag{45}$$

for all k. □

We see that under the foregoing assumptions, *no* other sampling process has a variance which is lower than that of periodic sampling. As just one example, by taking F to be exponential in $\hat{p}_1$ and inter-sample times to be independent, we learn that Poisson sampling yields a higher variance than periodic. However, the result is much more powerful than this. It shows that, if $R(\tau)$ is convex, no kind of train or other structure, no matter how sophisticated, can do better than periodic.

Unfortunately, periodic sampling has a disadvantage already discussed: it is not mixing, which makes it vulnerable to phase locking effects. Assuming that $R(\tau)$ is convex, we now determine sampling schemes that offer the best of both worlds: mixing to guarantee asymptotic consistency, but with variance close to that offered by periodic sampling.

For this, we will consider sampling using renewal processes with inter-probe times that are Gamma-distributed, namely with density

$$\Gamma_{\alpha,\lambda}(x) = \frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{\alpha-1}e^{-\lambda x}, \tag{46}$$

on $x > 0$, where $\Gamma(\cdot)$ is the familiar Gamma function. Its mean is $\mu^{-1} = \alpha/\lambda$ and its variance $\sigma^2 = \alpha/\lambda^2$. Gamma laws are well known to be stable with respect to the shape parameter α, that is, if $\{T_i \sim \Gamma_{\alpha_i,\lambda}\}$ are independent, then $\sum_i T_i \sim \Gamma_{\sum_i \alpha_i,\lambda}$. The exponential law corresponds to the 1-parameter sub-family $\Gamma_{1,\lambda}$. Another special sub-family are distributions with the Erlang law. These have only integral shape values.

We will need one more technical result regarding Gamma laws, the proof of which we leave to the Appendix.

Lemma 7 *Let $T \sim \Gamma_{\alpha,\lambda}$, $Z \sim \Gamma_{\beta,\lambda}$ be independent, and set $Y = T + Z$. Then $C = \mathbb{E}[T|Y] = \alpha Y/(\alpha+\beta)$ has density $\Gamma_{\alpha+\beta,(\alpha+\beta)\lambda/\alpha}$, with mean $\mathbb{E}[C] = a/\lambda = \mathbb{E}[T]$.*

We can now prove the following:

Theorem 4 *The family of renewal sampling processes $G(\beta)$, parameterized by $\beta > 0$, with inter-sample time density $\Gamma_{\beta,\beta\lambda}(x)$, provides, at constant mean sampling rate λ, sampling variance for $\hat{p}_1$ that monotonically decreases with β. The variance is larger (equal or smaller) than Poisson sampling as β is smaller (resp. equal or larger) than 1, and tends to that of periodic sampling in the limit $\beta \to \infty$.*

Proof We assume an underlying probability space on which the family of inter-sample variables are defined for each $\beta > 0$. Equation (43) holds for each inter-sample law $G(\beta)$. As the means for each are equal to $\mu = \beta/(\beta\lambda) = 1/\lambda$, proving the variance result reduces to showing that, for each $k > 0$, $\int R(\tau) f_{k,1}(d\tau) \geq \int R(\tau) f_{k,2}(d\tau)$ for any β values β_1, β_2 satisfying $\beta_2 > \beta_1$, where $f_{k,i}$ is the density of the sum $T_{k,i}$ of k inter-sample times, each with law $G(\beta_i)$. We can apply Jensen's inequality to show that

$$\mathbb{E}\big[\mathbb{E}\big[R(T_{k,1})\big|Y_{k,1}\big]\big] \geq \mathbb{E}\big[R\big(\mathbb{E}[T_{k,1}|Y_{k,1}]\big)\big] = \mathbb{E}\big[R(T_{k,2})\big] = \int R(\tau) f_{k,2}(d\tau)$$

where to show $\mathbb{E}[T_{k,1}|Y_{k,1}] = T_{k,2}$, we identified $(T, Y, \alpha, \beta, \lambda)$ with

$$\big(T_{k,1}, Y_{k,1}, k\beta_1, k(\beta_2 - \beta_1), \beta_1\lambda\big)$$

and used Lemma 7. Since this holds for any β_1, β_2 with $\beta_2 > \beta_1$, we have monotonicity of the variance in β. As β tends to infinity, there is weak convergence of

$\Gamma_{\beta,\beta\lambda}(x)(dx)$ to a Dirac measure at $1/\lambda$, as is easily seen using Laplace transforms. Since the function R is convex, it is continuous, and as it is also bounded (as a second-order process), the property

$$\lim_{\beta\to\infty}\int R(x)\Gamma_{\beta,\beta\lambda}(x)(dx)=\int R(x)\delta_{1/\lambda}(dx)$$

follows from the very definition of weak convergence. This shows that the limit of the variances of the Gamma renewal estimators is that of the deterministic probe case, namely the optimal variance. □

This result provides a family of sampling processes with the desired properties. By selecting $\beta > 1$, we can ensure lower (more precisely, no higher) variance than Poisson sampling. By selecting β large, we obtain sampling variance close to the lowest possible, whilst still using a mixing process. The important point is that the parameter β can be used to continuously tune for any desired trade-off, and to set the sampling variance arbitrarily close to the optimal case.

There is therefore a need to better understand what classes of queueing systems/networks lead to second-order state processes enjoying the above convexity property beyond the few classes quoted below.

5.2.1 Known convex examples

A natural question is, how likely is it that networks of interest satisfy the convexity property for delay and/or loss? There are simple systems for which exact results are known. For example, Ott [30] showed that convexity holds for the virtual work process (equal to the delay of probes with $x=0$) of the $M/G/1$ queue.

We now show that the loss process $I(t)$ of the $M/M/1/B$ queue (namely the indicator function that the number of customers is B, i.e. the set of periods where arriving packets are lost) has a convex auto-covariance function. Denote by λ and μ the arrival and the service rates and by $\rho=\lambda/\mu$ the load factor. From [40] (p. 13, Theorem 1), the probability that the number of customers in the queue is B at time t, given that it is B at time 0, is

$$P_{B,B}(t)=\frac{1-\rho}{1-\rho^{B+1}}\rho^B+\frac{2}{B+1}\sum_{j=1}^{B}\frac{\exp(-(\lambda+\mu)t+2t\sqrt{\lambda\mu}\cos(\pi j/(B+1))}{1-2\sqrt{\rho}\cos(\pi j/(B+1))+\rho}$$
$$\times\big(\sin\big(Bj\pi/(B+1)\big)-\sqrt{\rho}\sin(j\pi)\big)^2 \tag{47}$$

in the case when $\rho\neq 1$, and

$$P_{B,B}(t)=\frac{1}{1+B}+\frac{1}{B+1}\sum_{j=1}^{B}\frac{\exp(-(2\lambda)t+2\lambda t\cos(\pi j/(B+1))}{1-\cos(\pi j/(B+1))}$$
$$\times\big(\sin\big(Bj\pi/(B+1)\big)-\sin(j\pi)\big)^2 \tag{48}$$

in the case $\rho = 1$. In both cases, the auto-covariance function of $I_x(t)$, which is equal to $\pi(B)P_{B,B}(t)$ (with $\pi(B)$ the stationary probability that the queue has B customers) is a convex combination of convex decreasing functions of t and is hence itself convex and decreasing in t.

5.3 Maximum likelihood

Consider some network with a non-intrusive probing process N where the unknown parameters are obtained by some maximal likelihood method. An example of such a system would be that of Sect. 4.2 when the sequence of end-to-end delays idealized as i.i.d. seen by intrusive probes considered there is replaced by the sequence $\{W(T_n)\}$, where $W(t)$ is the virtual end-to-end delay in the network at time t. Hence $W(T_n)$ is the end-to-end delay seen by the nth (stealthy) probe. Here, $\{W(t)\}$ is a continuous time Markov chain and if N is an independent renewal process, then the sequence $\{W(T_n)\}$ is Markov. If one knows the transition kernel P_t of the continuous-time Markov chain $\{W(t)\}$, then one can compute the likelihood function associated with the samples $W(T_n)$, $1 \le n \le m$, through a formula that involves P_t and the stationary law of $\{W(t)\}$. Here are a few open problems within this setting:

- What renewal point processes are asymptotically efficient within this setting? We conjecture that if $\{W(t)\}$ is mixing, then all renewal point processes are asymptotically efficient.
- For m fixed, what renewal point process gives the MLE with the smallest variance among the set of all renewal point processes with intensity μ? Is the deterministic point process again optimal in terms of variance?

Acknowledgements The authors thank David McDonald and Patrick Thiran for their useful comments on this paper.

Appendix

6.1 Packet pairs in the $M/M/1$ queue

Consider an $M/M/1$ queue in steady state with the usual notation. Assume one sends to this system two additional customers at time 0 and $t > 0$ respectively, both with size x. Below we assume that t is small and that $x > t$. Let us denote by V_0 the system time of the first customer and by W_t that of the second. We are interested in the quantity $\mathbb{E}(V_0 W_t)$. Let S be an exponential random variable with parameter μ. Conditioned on the fact that the first customer finds an empty system, the latter is

$$(1-\lambda t)\big(x(2x-t)\big) + \lambda t \mathbb{E}\big(x(2x+S-t)\big) + o(t) = x(2x-t) + tx\rho + o(t).$$

Let $A(n)$ be the sum of n independent random variables, all exponential with parameter μ. Conditioned on the fact that the first customer finds n customers in the system, the quantity of interest is

$$\begin{aligned}&(1-\lambda t)\mathbb{E}\big[\big(A(n)+x\big)\big(A(n)+2x-t\big)\big]\\ &\quad + \lambda t \mathbb{E}\big[\big(A(n)+x\big)\big(A(n)+2x+S-t\big)\big] + o(t)\end{aligned}$$

$$= x(2x-t) + \mathbb{E}\big[A(n)^2\big] + (3x-t)\frac{n}{\mu} + \lambda t\left(\frac{n}{\mu}+x\right)\frac{1}{\mu} + o(t)$$
$$= x(2x-t) + n\frac{2}{\mu^2} + n(n-1)\frac{1}{\mu^2} + (3x-t)\frac{n}{\mu} + \lambda t\left(\frac{n}{\mu}+x\right)\frac{1}{\mu} + o(t).$$

Hence

$$\begin{aligned}\mathbb{E}(V_0 W_t) &= x(2x-t) + tx\rho + \frac{1}{\mu^2}\left(\frac{2\rho^2}{(1-\rho)^2} + \frac{\rho}{1-\rho}\right)\\ &\quad + \frac{\rho}{1-\rho}\left(\frac{3x-t}{\mu} + \frac{1}{\mu^2} + \rho t\frac{1}{\mu}\right) + o(t)\\ &= \mathbb{E}(V_0 W_0) - t\left((1-\rho)x + \rho\frac{1}{\mu}\right),\end{aligned}$$

with

$$\mathbb{E}(V_0 W_0) = 2x^2 + \frac{1}{\mu^2}\left(\frac{2\rho^2}{(1-\rho)^2} + \frac{\rho}{1-\rho}\right) + \frac{\rho}{1-\rho}\left(\frac{3x}{\mu} + \frac{1}{\mu^2}\right).$$

6.2 Realism of the Kelly model

This section discusses quickly the divergence between a Kelly model and a real IP backbone network. We refer to [22] for the impact of these differences on the estimators.

An instance of Internet path with two routers is depicted in Fig. 3. Actual routers may follow complex scheduling disciplines, and real packets experience delays on the incoming side and contention across the backplane, in addition to the output buffer queueing that the commonly used FIFO model nominally represents. However, it was shown in [20] that an output port in an IP router behaves like a single server FIFO queue. Traversing an IP network on the path from Source to Destination can hence be represented as some deterministic propagation delay, plus a random delay corresponding to traversing a series of single server FIFO queues with cross-traffic as exemplified in this figure.

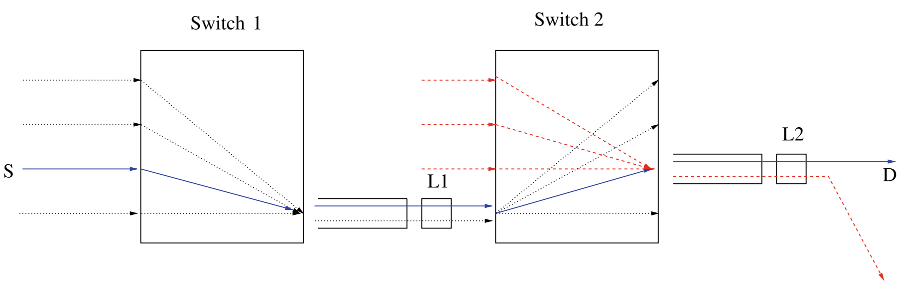

Fig. 3 Example of path with two routers with non-persistent cross traffic streams (*black dotted* and *red dashed*), whereas the probes (*blue*) pass end-to-end

The main restrictive assumption is that cross-traffic is Poisson, and that packets have an exponentially distributed sizes which are independently drawn at each hop. Here are a few comments on the matter.

It is of course well known that Internet traffic is not Poisson (e.g. [34]); for example, both the packet and TCP flow arrival processes exhibit long-range dependence. Although Poisson may nonetheless be a good assumption below some timescale (say 1 second [10, 19]) due to the 'noising' effect of multiplexing tens of thousands of largely independent flows.

It is well known (e.g. [11]) that the distribution of IP packets is strongly discrete, and can even be modeled as trimodal. For example, $S \in \{40, 576, 1500\}$ bytes, with probabilities $(0.5, 0.1, 0.4)$, captured its rough shape well in many cases. This is very far from exponential; however, its coefficient to mean ratio $\mathrm{Cov}[S] = \sqrt{\mathrm{Var}[S]}/\mathbb{E}[S] \approx 1.05$, which is very close to the 1 of the exponential case.

In real networks, packets have a size which, in terms of bytes (ignoring effects like changes in encapsulation), does not change as it traverses the network. In Kelly networks, packet sizes are modeled by service times which are chosen independently at each station. So, the parametric inversion methods based on Kelly networks will not work when 'persistent' traffic dominates (by persistent traffic, we mean cross-traffic which traverses more than one node of the monitored path).

6.3 Distribution of end-to-end delays in Kelly networks

Let us now consider the system when a tagged probe leaves the system. Since the queueing discipline is FIFO, the number of probes N in the system at that time is equal to the number of probes arrived during the time D the probe spent in the system. So denoting by $\phi(t)$ the density of D at $t \geq 0$, we get:

$$\mathbb{P}(N = k) = \int_0^\infty \phi(t)\mathbb{P}(N = k|D = t)\, dt = \int_0^\infty \phi(t)e^{-xt}\frac{(xt)^k}{k!}\, dt.$$

So the generating function $\psi_N(z)$ of the number of probes in the system at a probe departure epoch verifies:

$$\begin{aligned}\psi_N(z) &= \sum_{k\geq 0} z^k \mathbb{P}(N = k) = \sum_{k\geq 0}\int_0^\infty \phi(t)e^{-xt}\frac{(xtz)^k}{k!}\, dt\\ &= \int_0^\infty \phi(t)e^{-x(1-z)t}\, dt = \widehat{D}\big(x(1-z)\big),\end{aligned}$$

where $\widehat{D}(z)$ is the Laplace transform of D. Hence, setting $s = x(1-z)$ the Laplace transform of the end-to-end delay D is

$$\widehat{D}(s) = \psi_N\left(1 - \frac{s}{x}\right) = \prod_{j=1}^{K} \frac{\gamma_j - x}{\gamma_j - x + s}, \tag{49}$$

where we used the fact that ψ_N coincides with the steady-state distribution of the number of probes in the system (7), so that $\psi_N(z)$ is given by (8). Note that this is a function of residual capacities, but not the service or arrival rates, showing that only the γ_j are accessible by this technique.

Note that (49) is the product of the Laplace transform of exponential variables of parameters $\gamma_j - x$. By injectivity of the Laplace transform of random variables admitting a density, this proves that the end-to-end delay of probes is the sum of independent exponential random variables of parameters $\gamma_j - x$. The mean value is hence $\overline{D} = \sum_j \frac{1}{\gamma_j - x}$. Using the Laplace inversion formula and the residue theorem, and setting $\gamma_i' = \gamma_i - x$,

$$\phi(t) = \frac{1}{2\pi i} \int_{\alpha - i\cdot\infty}^{\alpha + i\cdot\infty} e^{st} \widehat{D}(s)\, ds = \sum \mathrm{Res}\left(e^{st} \prod_{i=1}^{K} \frac{\gamma_i'}{\gamma_i' + s} \right),$$

so that using $\alpha = 0$ and then the curve going from $-i\infty$ to $i\infty$ and back on a half-circle of infinite radius in the left half-plane, we get (14).

6.4 Proofs of the lemmas on E–M

6.4.1 Proof of Lemma 4

We have

$$\phi_{\theta_1,\theta_2}(\mathbf{l}, \mathbf{d}) = \phi_{\theta_1,\theta_2}(\mathbf{l}|\mathbf{d}) \phi_{\theta_1,\theta_2}(\mathbf{d}).$$

Hence

$$Q_{\mathbf{d}}(\theta_1, \theta_2 | \gamma_1, \gamma_2) = \mathbb{E}\big[\log\big(\phi_{\theta_1,\theta_2}(\mathbf{l}|\mathbf{d})\big)\big] + \log\big(f_{\mathbf{d}}(\theta_1, \theta_2)\big) \tag{50}$$

where the expectation (here and below) is w.r.t. the density

$$\phi_{\gamma_1,\gamma_2}(\mathbf{l}|\mathbf{d})$$

defined above. The log-likelihood function verifies:

$$\log f_{\mathbf{d}}(\theta_1, \theta_2) = Q_{\mathbf{d}}(\theta_1, \theta_2 | \gamma_1, \gamma_2) - \mathbb{E}\big[\log\big(\phi_{\theta_1,\theta_2}(l_1, \ldots, l_n | \mathbf{d})\big)\big]. \tag{51}$$

Using Jensen's inequality, it is easy to see that, for all (γ_1, γ_2) and (θ_1, θ_2),

$$\mathbb{E}\big[\log\big(\phi_{\theta_1,\theta_2}(\mathbf{l}|\mathbf{d})\big)\big] \le \mathbb{E}\big[\log\big(\phi_{\gamma_1,\gamma_2}(\mathbf{l}|\mathbf{d})\big)\big]$$

with equality iff

$$\phi_{\theta_1,\theta_2}(\mathbf{l}|\mathbf{d}) = \phi_{\gamma_1,\gamma_2}(\mathbf{l}|\mathbf{d}),$$

almost everywhere (this is in fact the information inequality). Using this last equality, (51) and (25), we get:

$$\begin{aligned} &\log f_{\mathbf{d}}\big(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}\big) - \log f_{\mathbf{d}}\big(\gamma_1^{(k)}, \gamma_2^{(k)}\big) \\ &\quad = Q_{\mathbf{d}}\big(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}\big) - Q_{\mathbf{d}}\big(\gamma_1^{(k)}, \gamma_2^{(k)}\big) \end{aligned}$$

$$- \mathbb{E}\big[\log\big(\phi_{\gamma_1^{(k+1)},\gamma_2^{(k+1)}}(\mathbf{l}|\mathbf{d})\big)\big] + \mathbb{E}\big[\log\big(\phi_{\gamma_1^{(k)},\gamma_2^{(k)}}(\mathbf{l}|\mathbf{d})\big)\big]. \tag{52}$$

The sum of the first two terms is positive due to optimality. The sum of the last two terms is positive as seen above. Hence

$$\log f_{\mathbf{d}}\big(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}\big) - \log f_{\mathbf{d}}\big(\gamma_1^{(k)}, \gamma_2^{(k)}\big) \geq 0, \tag{53}$$

with equality iff

$$\phi_{\gamma_1^{(k+1)},\gamma_2^{(k+1)}}(\mathbf{l}|\mathbf{d}) = \phi_{\gamma_1^{(k)},\gamma_2^{(k)}}(\mathbf{l}|\mathbf{d})$$

almost everywhere, and

$$Q_{\mathbf{d}}\big(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}|\gamma_1^{(k)}, \gamma_2^{(k)}\big) = Q_{\mathbf{d}}\big(\gamma_1^{(k)}, \gamma_2^{(k)}|\gamma_1^{(k)}, \gamma_2^{(k)}\big).$$

According to (53), the sequence $\log f_{\mathbf{d}}(\gamma_1^{(k)}, \gamma_2^{(k)}|d_1, \ldots, d_n)$ increases, and it is bounded by a constant. Therefore it converges.

6.4.2 Proof of Lemma 5

The proof relies on the following lemmas, which will be proved at the end of this section. Lemma 8 is technical and corresponds to a classical property of limit points of a sequence. It will be used only to prove Lemma 9, which will be the basic block of the result.

Lemma 8 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence with values in $\mathbb{R}$, s.t. $(x_{n+1} - x_n)$ converges to zero. Assume that a and b are both limit points of (x_n). Then every point c in $[a, b]$ is also a limit point of (x_n).*

Lemma 9 *Let (x_n) be a sequence with values in $\mathbb{R}$ and f a continuous function from $\mathbb{R}$ to $\mathbb{R}$. Assume that the sequence $(f(x_n))_{n\in\mathbb{N}}$ is convergent, and that (x_n) is bounded. Assume further that the following relation holds*:

$$f(x_{k+1}) - f(x_k) \geq g(x_{k+1} - x_k), \tag{54}$$

where $g(\cdot)$ is a positive continuous function, null at and only at zero. Then the sequence (x_n) is also convergent.

First, using (31), let express $\gamma_2^{(k)}$ as a function of $h(\cdot)$ of $\gamma_1^{(k)}$), where

$$h(x) = \left(\frac{\sum_{i=1}^n d_i}{n} - \frac{1}{x}\right)^{-1}.$$

Let us evaluate

$$\Delta_k = Q_{\mathbf{d}}\big(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}\big|\gamma_1^{(k)}, \gamma_2^{(k)}\big) - Q_{\mathbf{d}}\big(\gamma_1^{(k)}, \gamma_2^{(k)}\big|\gamma_1^{(k)}, \gamma_2^{(k)}\big).$$

Using (30) we get

$$\Delta_k = n\log\gamma_1^{(k+1)} + n\log\gamma_2^{(k+1)} - n\log\gamma_1^{(k)} - n\log\gamma_2^{(k)}$$

$$
\begin{aligned}
&-\gamma_1^{(k+1)}\sum d_i+\gamma_1^{(k)}\sum d_i\\
&-n\frac{\gamma_2^{(k+1)}-\gamma_1^{(k+1)}}{\gamma_2^{(k)}-\gamma_1^{(k)}}+n\frac{\gamma_2^{(k)}-\gamma_1^{(k)}}{\gamma_2^{(k)}-\gamma_1^{(k)}}\\
&+\left(\left(\gamma_2^{(k+1)}-\gamma_1^{(k+1)}\right)-\left(\gamma_2^{(k)}-\gamma_1^{(k)}\right)\right)\sum_i\frac{d_i}{e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}-1}.
\end{aligned}
$$

Using optimality (27) to re-express the sum $\sum_i \frac{d_i}{e^{(\gamma_2^{(k)}-\gamma_1^{(k)})d_i}-1}$ in terms of $\gamma_1^{(k)}, \gamma_2^{(k)}$ and $\gamma_2^{(k+1)}$, and using (31) to re-express the sum $\sum_i d_i$ in terms of $\gamma_1^{(k+1)}$ and $\gamma_2^{(k+1)}$, direct calculations reduce this to

$$
\Delta_k=ng\left(\frac{\gamma_1^{(k)}}{\gamma_1^{(k+1)}}\right)+ng\left(\frac{\gamma_2^{(k)}}{\gamma_2^{(k+1)}}\right)\geq ng\left(\frac{\gamma_1^{(k)}}{\gamma_1^{(k+1)}}\right),
$$

with $g(x)=x-1-\log(x)$.

Let $\mu_1^{(k)}=\log(\gamma_1^{(k)})$. Hence

$$
\Delta_k\geq g^*\left(\mu_1^{(k)}-\mu_1^{(k+1)}\right) \tag{55}
$$

with $g^*(x)=n(e^x-1-x)$. The last function is continuous, null at 0 and strictly positive elsewhere.

Let define now $f^*(x)=\log f_{\mathbf{d}}(e^x,h(e^x))$. We can therefore rename the sequence $\log f_{\mathbf{d}}(\gamma_1^{(k)},\gamma_2^{(k)})$ as $f^*(\mu_1^{(k)})$.

Lemma 4 shows that the sequence $f^*(\mu_1^{(k)})$ is convergent. The sequence $(\mu_1^{(k)})$ can be bounded by construction (see the proof of Theorem 2). Finally, (55) and (52) show that $f^*(\mu_1^{(k+1)})-f^*(\mu_1^{(k)})\geq g^*(\mu_1^{(k)}-\mu_1^{(k+1)})$. h and f^* are continuous at any point greater than $\log\frac{n}{\sum d_i}$, which will be the case after the first iteration. Therefore, Lemma 9 can be applied, the sequences $(\mu_1^{(k)})$ and $(\gamma_1^{(k)})=(e^{\mu_1^{(k)}})$ converge. As h is a continuous function, the sequence $(\gamma_2^{(k)})=(h(\gamma_2^{(k)}))$ is also convergent, and this will be the case for the sequence $(\gamma_1^{(k)},\gamma_2^{(k)})$.

We now prove that the limit is a solution of the likelihood equation. At any fixed point we have $\gamma_1^{(k+1)}=\gamma_1^{(k)}=\gamma_1^*$ and $\gamma_2^{(k+1)}=\gamma_2^{(k)}=\gamma_2^*$. Therefore, using (26) and (27):

$$
\frac{\gamma_2^*}{(\gamma_2^*-\gamma_1^*)\gamma_1^*}=\frac{1}{n}\sum_{i=1}^n\frac{d_ie^{(\gamma_2^*-\gamma_1^*)d_i}}{e^{(\gamma_2^*-\gamma_1^*)d_i}-1}
$$

and

$$
\frac{\gamma_1^*}{(\gamma_2^*-\gamma_1^*)\gamma_2^*}=\frac{1}{n}\sum_{i=1}^n\frac{d_i}{e^{(\gamma_2^*-\gamma_1^*)d_i}-1},
$$

the same equations as the likelihood equation, which means that any fixed point of the E–M algorithm is also a solution of the likelihood equation.

Proof of Lemma 8 Let c be a point in $[a, b]$, and let construct a sub-sequence of (x_n) that converges toward c. By definition, a and b are limit points of (x_n), hence we can assume that $c \neq a$ and $c \neq b$.

Let $\epsilon = \min(\frac{c-a}{2}, \frac{b-c}{2})$ be a positive number. $(x_{n+1} - x_n)$ converges toward zero. Hence, $\forall k, \exists N_k$ s.t. $\forall j > N_k, x_{j+1} - x_j < \frac{\epsilon}{k}$. By definition of a limit point, $\exists i_0 \leq N_1$, s.t. $x_{i_0} \in]a-\epsilon, a+\epsilon[$. Similarly, $\exists j_0 > i_0$, s.t. $x_{j_0} \in]b-\epsilon, b+\epsilon[$. Recursively, we can construct two integer sequences (i_k) and (j_k), such that $\forall k$, $N_{k+1} \leq i_k < j_k < i_{k+1}$, $i_k \in]a-\epsilon, a+\epsilon[$ and $j_k \in]b-\epsilon, b+\epsilon[$.

We are now in position to conclude. For all k, we have that $x_{i_k} < a + \epsilon \leq c - \frac{\epsilon}{k} < c + \frac{\epsilon}{k} \leq b - \epsilon < x_{j_k}$. Furthermore, $i_k < j_k$, and $\forall n > i_k$, $(x_{n+1} - x_n) \leq \frac{\epsilon}{k}$. This is enough to conclude that there exists $i_k < \phi(k) < j_k$ such that $x_{\phi(k)} \in]c - \frac{\epsilon}{k}, c + \frac{\epsilon}{k}[$. Since $j_k < i_{k+1}$, the function $\phi(.)$ is strictly growing, and $(x_{\phi(n)})$ is a sub-sequence of (x_n) convergent toward c. □

Proof of Lemma 9 By assumption, the sequence (x_n) is bounded. Therefore, it converges if and only if it admits one unique limit point.

Assume, by contradiction, that there are two distinct limit points a and b, with $a < b$. The sequence $f(x_n)$ is convergent, therefore $(f(x_{n+1}) - f(x_n))$ converges toward zero. Using (54), we get that $g(x_{n+1} - x_n)$ is convergent toward zero. By contradiction, if $(x_{n+1} - x_n)$ admits one limit point $c \neq 0$, then the sequence $g(x_{n+1} - x_n)$ admits $g(c) > 0$ as a limit point, which contradicts to the fact that it converges to 0. Hence, $(x_{n+1} - x_n)$ admits no non-zero limit point, and as it is bounded, converges to 0.

Using Lemma 8, we get that $\forall c \in [a, b]$, c is a limit point of (x_n), and hence $f(c)$ is a limit point of $f(x_n)$. As $f(x_n)$ converges toward l, it admits one unique limit point, and $\forall c \in [a, b]$, $f(c) = f(a) = f(b) = l$. Let $\epsilon = \frac{b-a}{3}$ be a positive number, and let now N be such that $\forall n \geq N$, $|x_{n+1} - x_n| < \epsilon$. As a and b are limit points of (x_n), there exist $n_1 \geq N$ and $n_2 \geq n_1$ such that $|x_{n_1} - a| < \epsilon$ and $|x_{n_2} - b| < \epsilon$. Then $\exists n_3$ s.t. $n_1 \leq n_3 < n_3 + 1 \leq n_2$, $x_{n_3} \neq x_{n_3+1}$, $x_{n_3} \in]a, b[$ and $x_{n_3+1} \in]a, b[$. On the one hand, (54) leads to $f(x_{n_3+1}) > f(x_{n_3})$. On the other hand, $f(x_{n_3}) = f(x_{n_3+1}) = l$. We get a contradiction. □

6.5 Proof of Lemma 6

For the three-routers case, for all $k \geq 0$, (25) is equivalent to

$$\begin{aligned} &Q(\theta_1, \theta_2, \theta_3 | \gamma_1, \gamma_2, \gamma_3) \\ &\quad = \sum_{i=1}^{n} \alpha(d_i) \int_{l_1=0}^{d_i} \int_{l_2=0}^{d_i - l_1} \big(\log(\theta_1\theta_2\theta_3) - \theta_3 d_i + (\theta_3 - \theta_1) l_1 + (\theta_3 - \theta_2) l_2\big) \\ &\qquad \times e^{(\gamma_3-\gamma_1)l_1} e^{(\gamma_3-\gamma_2)l_2}\, dl_2\, dl_1 \end{aligned} \tag{56}$$

with α defined in (33). We have

$$\int_{l_1=0}^{d} \int_{l_2=0}^{d-l_1} (a + b l_1 + c l_2) e^{\alpha l_1} e^{\beta l_2}\, dl_2\, dl_1 = a c_a + b c_b + d c_d$$

with

$$c_a = \frac{\beta(e^{\alpha d} - 1) - \alpha(e^{\beta d} - 1)}{\alpha\beta(\alpha - \beta)},$$

$$c_b = \frac{\alpha\beta(\alpha - \beta)de^{\alpha d} - \beta(2\alpha - \beta)e^{\alpha d} + \alpha^2 e^{\beta d} - (\alpha - \beta)^2}{\alpha^2\beta(\alpha - \beta)^2},$$

$$c_c = \frac{\alpha(\alpha - 2\beta)e^{\beta d} - \alpha\beta(\alpha - \beta)de^{\beta d} + \beta^2 e^{\alpha d} - (\alpha - \beta)^2}{\alpha\beta^2(\alpha - \beta)^2}.$$

In order to evaluate (56) we have to take $\alpha = \gamma_3 - \gamma_1$, $\beta = \gamma_3 - \gamma_2$, $a = \log(\theta_1\theta_2\theta_3) - \theta_3 d_i$, $b = (\theta_3 - \theta_1)$ and $c = (\theta_3 - \theta_2)$ (note that $\alpha - \beta = \gamma_2 - \gamma_1$). In addition,

$$\alpha(d_i) = \frac{\alpha\beta(\alpha - \beta)}{\beta e^{\alpha d} - \alpha e^{\beta d} + \beta - \alpha} = \frac{1}{c_a}.$$

Finally

$$\begin{aligned} &Q(\theta_1, \theta_2, \theta_3 | \gamma_1, \gamma_2, \gamma_3) \\ &= \sum_{i=1}^{n} \alpha(d_i)\big[c_a \log(\theta_1\theta_2\theta_3) - c_b\theta_1 - c_c\theta_2 + (c_b + c_c - d_i c_a)\theta_3\big] \\ &= n\log(\theta_1\theta_2\theta_3) + \sum_{i=1}^{n}\left(\frac{c_b + c_c}{c_a} - d_i\right)\theta_3 - \frac{c_b}{c_a}\theta_1 - \frac{c_c}{c_a}\theta_2, \end{aligned}$$

and therefore

$$\frac{\partial Q(\theta_1, \theta_2, \theta_3 | \gamma_1, \gamma_2, \gamma_3)}{\partial\theta_1} = \frac{n}{\theta_1} - \sum_{i=1}^{n}\frac{c_b}{c_a}.$$

The expressions given in the lemma are then directly obtained from

$$\frac{\partial Q(\gamma_1^{(k+1)}, \gamma_2^{(k+1)}, \gamma_3^{(k+1)} | \gamma_1^{(k)}, \gamma_2^{(k)}, \gamma_3^{(k)})}{\partial\gamma_1^{(k+1)}} = 0$$

and other relations of the same type.

6.6 Proof of Lemma 7

Let $T \sim \Gamma_{\alpha,\lambda}$, $Z \sim \Gamma_{\beta,\lambda}$ be independent, and set $Y = T + Z$. Then $C = \mathbb{E}[T|Y] = \alpha Y/(\alpha + \beta)$ has density $\Gamma_{\alpha+\beta,(\alpha+\beta)\lambda/\alpha}$, with mean $\mathbb{E}[C] = a/\lambda = \mathbb{E}[T]$.

Proof From the scaling property of Gamma, $Y \sim \Gamma_{\alpha+\beta,\lambda}$. Since T and Z are independent, the density of $(T|Y=y)$ is

$$\begin{aligned} P(T=x|Y=y) &= \frac{P(T=x, Y=y)}{P(Y=y)} = \frac{P(T=x, Z=y-x)}{P(Y=y)} \\ &= \frac{\Gamma_{\alpha,\lambda}(x)\Gamma_{\beta,\lambda}(y-x)}{\Gamma_{\alpha+\beta,\lambda}(y)} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(y-x)^{\beta-1} y^{1-(\alpha+\beta)}. \end{aligned}$$

Recall the *Beta function* $B(x,y) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. The required conditional expectation is given by

$$\begin{aligned} \mathbb{E}[T|Y=y] &= \frac{y^{1-(\alpha+\beta)}}{B(\alpha,\beta)} \int_0^y x^{\alpha}(y-x)^{\beta-1}\,dx \\ &= \frac{y^{1-(\alpha+\beta)}}{B(\alpha,\beta)} y^{\alpha+\beta} B(\alpha+1,\beta) \\ &= \frac{\alpha y}{\alpha+\beta} \end{aligned} \tag{57}$$

using the integral identity 3.191(1) from [17]. Now viewing y as a sample of Y, we have $C = \mathbb{E}[T|Y] = \alpha Y/(\alpha+\beta)$, which is Gamma as stated by the scaling property. □

References

1. Alouf, S., Nain, P., Towsley, D.F.: Inferring network characteristics via moment-based estimators. In: INFOCOM, pp. 1045–1054 (2001)
2. Arya, V., Duffield, N., Veitch, D.: Multicast inference of temporal loss characteristics. In: IFIP Performance 2007, Special Issue, Performance Evaluation, Cologne, Germany, 2–5 October 2007. Elsevier, Amsterdam (2007)
3. Asmussen, S., Nerman, O., Olsson, M.: Fitting phase-type distributions via the em algorithm. Scand. J. Stat. **23**(4), 419–441 (1996)
4. Baccelli, F., Bremaud, P.: Elements of Queueing Theory Applications of Mathematics, 2nd edn. Springer, Berlin (2003)
5. Baccelli, F., Machiraju, S., Veitch, D., Bolot, J.: The role of PASTA in network measurement. Comput. Commun. Rev. **36**(4), 231–242 (2006). Proceedings of ACM SIGCOMM 2006
6. Baccelli, F., Machiraju, S., Veitch, D., Bolot, J.: On optimal probing for delay and loss measurement. In: Proc. ACM SIGCOMM Internet Measurement Conf., San Diego, California, 23–26 October 2007, pp. 291–302 (2007)
7. Baccelli, F., Massey, W., Towsley, D.: Acyclic fork-join queuing networks. J. ACM **36**(3), 615–642 (1989)
8. Bolot, J.C.: Characterizing end-to-end packet delay and loss in the Internet. J. High-Speed Netw. **2**(3), 305–323 (1993)
9. Caceres, R., Duffield, N., Horowitz, J., Towsley, D.: Multicast-based inference of network-internal loss characteristics. IEEE Trans. Inf. Theory **45**, 2462–2480 (1999)
10. Cao, J., Cleveland, W.S., Lin, D., Sun, D.X.: On the nonstationarity of Internet traffic. In: SIGMETRICS'01: Proc. ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, pp. 102–112. ACM, New York (2001)

11. Claffy, K., Miller, G., Thompson, K.: The nature of the beast: recent traffic measurements from an Internet backbone. In: INET'98, Geneva, Switzerland, 21–24 July 1998. Internet Society (1998)
12. Coates, M., Nowak, R.: Network tomography for internal delay estimation. In: Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), Salt Lake City, Utah, May 2001
13. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (1), 1–38 (1977)
14. Duffield, N., Horowitz, J., Presti, F.L., Towsley, D.: Multicast topology inference from measured end-to-end loss. IEEE Trans. Inf. Theory **48**(1), 26–45 (2002)
15. Glynn, P., Sigman, K.: Independent sampling of a stochastic process. Stoch. Process. Appl. **78**, 151–164 (1998)
16. Golub, G.H., Loan, C.V.: An analysis of the total least squares problem. Technical Report, Ithaca, NY, USA (1980)
17. Gradshteyn, L., Ryzhik, L.: Table of Integrals, Series and Products, 6th edn. Academic Press, San Diego (2000)
18. Grenouille. http://grenouille.com/cest_quoi.php
19. Hohn, N., Veitch, D., Abry, P.: Cluster processes, a natural language for network traffic. IEEE Trans. Signal Process. **51**(8), 2229–2244 (2003). Special issue "Signal Processing in Networking"
20. Hohn, N., Veitch, D., Papagiannaki, K., Diot, C.: Bridging router performance and queuing theory. In: Proc. ACM SIGMETRICS, New York, 12–16 June 2004, pp. 355–366 (2004)
21. Kapoor, R., Chen, L.-J., Lao, L., Gerla, M., Sanadidi, M.: CapProbe: a simple and accurate capacity estimation technique. In: Proc. ACM SIGCOMM, pp. 67–78. ACM, New York (2004)
22. Kauffmann, B., Baccelli, F., Veitch, D.: Towards multihop available bandwidth estimation. ACM SIGMETRICS Perform. Eval. Rev. **37**(2), 84 (2009)
23. Kelly, F.: Reversibility and Stochastic Networks. Wiley, New York (1979)
24. Lai, K., Baker, M.: Measuring link bandwidths using a deterministic model of packet delay. In: Proc. ACM SIGCOMM'00, Stockholm, September 2000, pp. 283–294 (2000)
25. Larson, R.: The queue inference engine: deducing queue statistics from transactional data. Manag. Sci. **36**(5), 586–60 (1990)
26. Liu, X., Ravindran, K., Liu, B., Loguinov, D.: Single-hop probing asymptotics in available bandwidth estimation: sample-path analysis. In: Proc. ACM Internet Measurement Conf., October 2004
27. Liu, X., Ravindran, K., Loguinov, D.: Multi-hop probing asymptotics in available bandwidth estimation: stochastic analysis. In: Proc. ACM/USENIX Internet Measurement Conf., October 2005
28. Machiraju, S., Veitch, D., Baccelli, F., Bolot, J.: Adding definition to active probing. ACM Comput. Commun. Rev. **37**(2), 17–28 (2007)
29. Novak, A., Taylor, P., Veitch, D.: The distribution of the number of arrivals in a subinterval of a busy period in a single server queue. Queueing Syst. **53**(3), 105–114 (2006)
30. Ott, T.: The covariance function of the virtual waiting time process in an $M/G/1$ queue. Adv. Appl. Probab. **9** (1997)
31. Parker, B.: Design of experiments for packet networks. Ph.D. Thesis, Queen Mary, University of London, School of Mathematical Sciences (2009)
32. Pásztor, A., Veitch, D.: On the scope of end-to-end probing methods. IEEE Commun. Lett. **6**(11), 509–511 (2002)
33. Paxson, V.: End-to-end Internet packet dynamics. IEEE/ACM Trans. Netw. **7**(3), 277–292 (1999)
34. Paxson, V., Floyd, S.: Wide-area traffic: the failure of Poisson modeling. IEEE/ACM Trans. Netw. **3**(3), 226–244 (1994)
35. Petersen, K.: Ergodic Theory. Cambridge University Press, Cambridge (1983)
36. Prasad, R.S., Murray, M., Dovrolis, C., Claffy, K.: Bandwidth estimation: metrics, measurement techniques, and tools. IEEE Netw. Mag. **17**(6), 27–35 (2003)
37. Schervish, M.: Theory of Statistics. Springer, New York (1995)
38. Sharma, V., Mazumdar, R.: Estimating traffic parameters in queueing systems with local information. Perform. Eval. **32**, 217–230 (1998)
39. Strauss, J., Katabi, D., Kaashoek, F.: A measurement study of available bandwidth estimation tools. In: Proc. ACM SIGCOMM Internet Measurement Conf., 2003
40. Takács, L.: Introduction to the Theory of Queues. Oxford University Press, New York (1962)